Introduction to Microwaves

The quality of the materials used in the manufacture of this book is governed by continued postwar shortages.

Introduction to
MICROWAVES

★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★

BY

SIMON RAMO, PH.D.
Director of Research, Radio Division,
Hughes Aircraft Company

First Edition
SIXTH IMPRESSION

★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★

McGraw-Hill Book Company, Inc.
NEW YORK AND LONDON
1945

INTRODUCTION TO MICROWAVES

PRINTED IN THE UNITED STATES OF AMERICA

THE MAPLE PRESS COMPANY, YORK, PA.

Preface

★ ★ ★ ★ ★ ★ ★ ★

"Microwaves" is a name popularly given to electricity when its frequency of alternation is in billions of cycles per second. This frequency ratio of millions when compared with power frequencies and of thousands when compared even with radio-broadcast frequencies makes microwaves fascinating and important. It also means, however, that for many whose acquaintance with electricity is with the lower frequency range only there is a need for further attention to fundamentals. It is hoped that at least for such readers this text will prove helpful. Physical pictures and viewpoints are described that have been found useful both to those doing active work in this relatively new branch of engineering and to others who desire merely a better appreciation of it.

Anyone who has spent the last several years investigating and trying to apply microwaves has been called upon to offer simple explanations of the often seemingly mysterious yet fundamental characteristics of extremely high frequency electricity. What makes waves go down the inside of a hollow conducting pipe at high frequency but not at low frequency? Do microwaves obey different laws than do low-frequency waves? How is it that high-frequency current can be drawn from electrodes of an electronic tube when the electron current in the tube does not even flow to those electrodes? May we still speak of impedance?

Such questions are often asked by commercial men, executives, and many engineers who specialize in other phases of electricity and have only a remote interest in microwaves. These people do not want a long, thorough explanation but they do desire a connecting link with their previous background. However, even engineers concerned with the most technical aspects of microwaves have a need for simple physical pictures and, indeed, will ask exactly the same questions during some period of their learning of the microwave art. Young engineers particularly, forced now to gain design experience rapidly, are often confused in the application of fundamentals. Their efficiency can often be materially increased by simple qualitative discussions that act as catalysts in the orderly arrangement of their knowledge.

Thus the thought that it might be helpful to write a nonmathematical discussion on concepts as an introduction to microwaves grew from observations of the apparent usefulness of such discussions carried on informally with the author's associates. The present book stems more directly, however, from notes taken during a lecture tour two years ago. Because of the general interest shown then, it seemed worth while to set down some important points in a more or less orderly fashion.

Dangers are recognized in presenting only physical pictures and conceptions unaccompanied by quantitative data and descriptions of apparatus. In the first place, one person's way of looking at a problem is not always sure to be helpful to another. Then, too, some may think there is an intention to discount the need for a complete treatment; they may feel that a complex field is being oversimplified. The present discussion is certainly not aimed to replace anything. It is merely a preface. For some, such an introduction may be sufficient. The book is most certainly a failure if it does not make clear the absolute

need for further study on the part of those who must do actual microwave engineering. The hope is that a number of important points may be cleared up in a few hours' reading and the way opened for clearer thought on some of the remaining phases. Some texts suitable for further study are listed in the appendix.

The author is indebted to J. R. Whinnery both for the numerous general discussions on microwave theory during the past several years that have added background for the text and for criticisms of the text material. Thanks are also expressed to W. C. White, who read the manuscript and made many suggestions as to exposition and choice of material.

SIMON RAMO.

SCHENECTADY, N. Y.
February, 1945.

Contents

* * * * * * * *

INTRODUCTION TO MICROWAVES

CHAPTER I

Similarities Exist between Microwaves and Lower Frequency Electricity

It is assumed that the reader's familiarity with electricity is limited to power frequencies or at least to the lower radio frequencies. This familiarity may be somewhat insecure at present, in view of the fairly recent appearance of apparently new phenomena like wave guides and resonant cavities, which we learn exist at ultra-high frequency; but the following discussion will attempt to raise the reader's understanding of basic principles of ultra-high frequency to the same level as his understanding of lower frequencies. It might be guessed that he can hardly acquire more knowledge about ultra-high-frequency concepts without adding to his appreciation of the lower frequencies. This is indeed the case, because from one point of view there is no basic difference between the extremely high and the very lowest frequency electricity.

It should serve the stated objective if it is made clear in what ways the high frequencies are the same, and in what other ways different from the lower frequencies. This chapter states in what ways electricity is the same over the whole frequency range. The word "states" is used because there will be no immediate attempt to justify the similarities, but only to list them; it will require much of the rest of the book to develop the physical pictures necessary for understanding this fundamental identity, or unity, of electricity over the entire power and radio spectrum.

Wave Action.—Thus it might be pointed out, as a beginning, that electricity, at any frequency, can be properly regarded as *wave* action. Whether the problem is one involving circuits, vacuum tubes, transformers, transmission lines, or antennas, the electrical phenomena can all be attributed, from one point of view, to a series of waves. That is what the physicists mean when they draw the spectrum of electromagnetic energy reproduced in Fig. 1. It has been provided with convenient markers to indicate which regions belong to the power field, which to the ordinary radiobroadcast range, and which to the ultra-high frequencies or microwaves. Beyond this point, the electro-

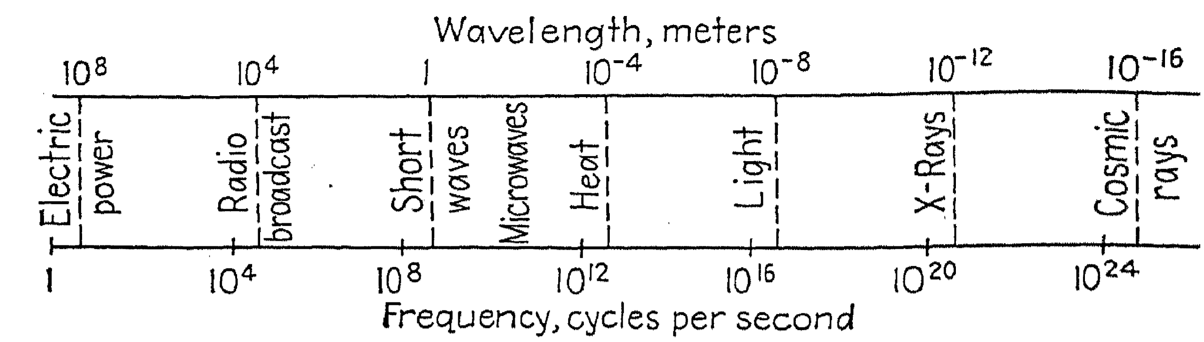

FIG. 1.—The electromagnetic spectrum.

magnetic waves are no longer called "ultra-high-frequency electricity," and, instead, the names "heat," "light," "X rays," and finally "cosmic rays" are used.

Basic Similarities.—Each region of the spectrum has its own descriptive language and its series of engineering applications. But this chapter is concerned with basic similarities; it seeks to get under the differences in language to seek identities, if any exist, in the laws of physics that apply to the various frequency bands. If this is done, it will be observed that the laws are the same at least over the range from zero frequency to well beyond the centimeter waves. Now it is necessary to qualify this statement somewhat and to interpret it carefully.

Maxwell, using what are now called the "laws of classical electricity and magnetism," was able to explain the chief characteristics of light, *i.e.*, its bending, diffraction, and

propagation. As a matter of fact, this electromagnetic theory of light furnished a real link between what were thought to be completely separate branches of physics—electricity and light. This association was in itself a good enough reason to draw out the spectrum of Fig. 1, suggesting that light waves are simply radio waves that happen to be short in wavelength. With this point of view, it may be said that heat waves lie between the light and centimeter-long radio waves; heat rays, and even X rays all exhibit properties of electromagnetic waves differing in their respective frequencies of oscillation. Considered as electromagnetic waves, they are characterized by the same velocity of propagation. The frequency specification leads immediately to a knowledge of how far in space the wave would travel in the time it would take two successive crests of the wave to pass a stationary observer (Fig. 2); *i.e.*, because the velocity is fixed, the wavelength, defined as the distance between these crests, is known when the frequency is given.

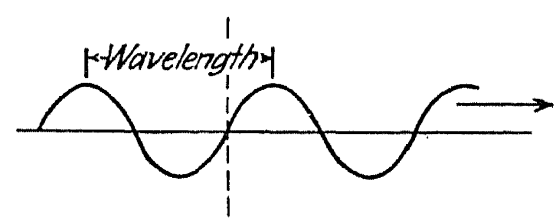

FIG. 2.—Velocity = frequency × wavelength.

As for the qualifications mentioned previously, the laws of classical electricity and magnetism do not explain all the characteristics of the phenomena whose labels appear on the spectrum. There are several reservations in the picture, but they need not concern us, with our rather limited objectives. Mainly they have to do with atomic and molecular phenomena. Electrons sometimes act like light waves or radio or heat rays of extremely short wavelength, but they have other characteristics as well. For instance, the very emission of electrons, as well as of X rays, heat, and light, involves complex physical laws that are different from, or at least not covered by, those classical electricity laws. As another example, the detailed nature of what goes on when electric current, even direct current, passes through a conductor involves atoms and molecules

and their relatively complicated and less familiar laws. These situations are not, however, a reason for withdrawing the statement that titles this chapter. What we want to do is to understand electricity as well at ultra-high frequency as at the lower frequencies. Thus are ruled out such things as the emission of electrons from a hot wire in a vacuum or the theory of matter and its properties like conductivity and permeability. We simply want to understand the basic over-all concepts such as current, charge, voltage, impedance, circuits, resonant cavities, and wave guiding. What takes place inside the atoms and molecules of a conductor to make Ohm's law a fact is a matter not explained by classical electricity and magnetism at any frequency. This classical theory, which serves the electrical engineer well, simply recites and accepts Ohm's law without further explanation.

Laws of Electricity for Whole Frequency Spectrum.—On this basis, the laws of electricity that will be of interest are truly the same laws over the whole frequency spectrum. That is perhaps the first thing to learn about ultra-high frequency. All apparent differences are differences in magnitude alone, not in fundamental laws.

Since there is no difference in the true fundamentals over the frequency range from direct current through the microwaves, then there is immediately a second important fact to learn about ultra-high frequency: If we find we do not understand its fundamentals, then it must be recognized that we have not yet fully understood the fundamentals of low frequency. Certainly, we may comprehend low frequency or medium frequency well enough to obtain at all times the correct answer to our problems. But that may simply mean that the magnitude of any possible error is too small to notice and is not to be taken as proof that the particular principles being applied are exactly correct.

Changes in Magnitude.—Let us continue to state the implications of the point of view that microwaves are not different from lower frequency electricity. If this is a

fact, only the change in the magnitudes of various effects as the frequency is increased need be considered. In the first place, it is agreed that electrons are still extremely small compared with the wavelengths of microwaves. This is to say that any direct effects of the discrete character of individual electrons will snobbishly continue to regard microwaves as next door to direct current. Then there are other effects. For instance, radiation of energy takes place in general from all electrical systems, circuits, transmission lines, antennas. It is completely negligible for systems, like low-frequency circuits, that are small compared with wavelength; it is large for antennas, for they are purposely made comparable in size to the wavelength. Radiation is simply an effect which grows in importance for a given configuration of electrical system driven at a fixed frequency as the physical size of the system grows. Or, as the frequency is increased on a given system, the radiation accordingly increases. The point is that radiation is predicted by the same laws as those existing over the entire frequency band of interest. There is no point of sudden shift in the basic principles; *i.e.*, the law is *not* the following: Below some point in frequency, radiation is impossible; above that frequency, radiation is described by a certain law which takes over authority.

Similarly, there is no fundamentally new effect in transmission-line or wave-guiding theory. With the use of a given physical size to wavelength ratio, any kind of wave-guiding system will be the same at any frequency. (Some factors, such as ohmic losses, may differ unless a proper scale factor is also applied to the constants of the material, such as conductivity; but this does not count as a basic difference between low and high frequency.)

Skin Effect.—A third example of a characteristic of electricity which is fundamentally the same over the frequency range but which varies enormously in magnitude of its importance is the skin effect. When carrying direct current, a long conductor has a uniform current distribu-

tion over its cross section. As the frequency increases, the ratio of the current that flows near the surface to the total current smoothly increases. When the wavelength becomes very small, the penetration of current becomes vanishingly small. The accent here is on the fact that the shifting of current density is a continuous one with frequency, and it begins to take place, unnoticed perhaps, as soon as any variation with time takes place in the current flow.

A mere statement of the fact that the fundamental laws are the same for low frequency and microwaves may be useful as a piece of information alone, but it does not yet provide the reader with physical pictures of ultra-high frequency. It is necessary to see the how and why of this statement. But even if the notion is granted, the reader should have questions. There must be common concepts that have been accepted because so many have worked so long at the lower frequencies—concepts that must be rejected or expanded as frequency increases. What are those concepts? What are the new ones, or how must the old ones be broadened? There may well be effects, interesting and important, that may be practical to attain with microwaves and that, though fundamentally possible at lower frequencies, have never been experienced for longer waves simply because the structure would have to be too huge for anyone to have contemplated building it.

CHAPTER II

Microwaves Are Very Different from Lower Frequency Electricity

One approach to science is to seek one all-inclusive law of nature. Then everything that can possibly happen is a special case of the same law. There is no fundamental difference between any two things, all apparent differences being simply a matter of magnitudes.

The engineer cannot be too happy with this approach. His situation is somewhat like that of the bass fiddler in the symphony orchestra who is told that, starting next week, he must occupy a seat in the first-violin section. To help him make the transition quickly, he is first of all given the assurance that there is no fundamental distinction between the two instruments, only differences in the magnitude of the effects. They are both stringed instruments, bowed or plucked, one hand for stopping the strings, the other for exciting them. Of course, in one instance the instrument is held horizontally in the arms of the player, and in the other the instrument is vertical with essentially all of its weight resting on the floor. Still, these things simply come about because of the differences in the magnitudes of the physical dimensions and weights of the fiddles.

Now the engineer is the man who makes, sells, strings, tunes, and plays the instruments. To him it is a basic distinction that the sizes, weights, variety of effects, and pitch of the instruments vary. To master either, he gradually comes upon and uses concepts which, though perhaps narrow in their application, enable him to think easily about the possibilities and applications of each narrow field. These useful concepts are often the results of excellent approximations (in the field in which they originated and for which they were intended).

Circuits at Low Frequencies.—At low frequencies there are certain good conceptions of circuits that may fail completely to describe similar effects at ultra-high frequencies. One of them is that losses take place only in ohmic dissipation; there are no losses through radiation, *i.e.*, by propagation of energy away from the circuit in electromagnetic waves. If radiation takes place, the circuit is conceived of as an antenna and not a circuit. Another conception is that current flows in the conductors of the circuit and there only. Still another is that an electric charge on the conducting surfaces of a circuit produces an electric field that changes in phase with the charge, and such capacitive fields, or the voltage drops they cause, are exactly 90 deg. out of time phase with the current flow. Similarly, a flow of current is thought of as being surrounded by a magnetic field in exact phase with the current, and the rate of change of the magnetic field gives rise to voltage drops in the conductor that are exactly 90 deg. out of phase with the current. There are many other common notions of circuits that are completely valid and useful when applied to certain regions of the frequency spectrum but, like those mentioned above, are applied with risk to the microwave region.

Low-frequency Transmission Line.—Even transmission-line theory at the lower frequencies has associated with it prevalent notions that are justified, strictly speaking, only in certain parts of the frequency spectrum. Thus it is common to think of a two-conductor transmission line as completely describable in terms of a single series impedance per unit length (the reactive part of which is an inductance computable from static field linkages) and a single admittance per unit length. There may be losses in each of these, ohmic losses, of course, but the narrow concept would have the energy dissipated only in such losses or in the load and none straying away by radiation.

This is not to say that radiation from lines, as well as any other factor mentioned so far, has been overlooked at

the lower frequencies, and that it remained for the ultra-high-frequency experts to come along and shatter the illusions. But these factors are most often too small and unimportant; it is natural, even efficient for most things, to develop concepts applicable only to the great majority of situations.

Microwave Band.—Figure 1 shows how the microwave band is located with respect to the ordinary radiobroadcasting band on one side and the heat and light waves on the other. If the relative spacing of these names on the spectrum is noted, it is difficult to refrain from asking why it should be expected that microwaves pick up more concepts from the power, or ordinary, radio bands than from the heat or light bands. Offhand, microwaves would be expected to act a little like each neighbor. And in connection with relatively narrow but highly useful concepts that are associated with each region, the great variety in the common concepts on each side of the microwaves should be noted. Thus in the electric power and radiobroadcast regions, the common concepts have to do with charges, currents, voltages, impedances, circuits, and I^2R losses. In heat we speak of temperature, calories, specific heats, and entropy. In light some of the common concepts are index of refraction, mirrors, lenses, prisms, and focal lengths. From these thoughts the impression is obtained that surely the microwave region can be best thought of in terms of many concepts. Some will resemble the common ones of lower frequency electricity. Others will be closely related to those of heat and light; and still others will perhaps be relatively narrow, peculiar to the microwave region alone. Of course, there will always be a liberal number of concepts that are the same over the whole range. These will be concepts that are not better replaced, from the standpoint of efficiency and understanding of any one band, by approximate measures applicable only to that band.

Common Microwave Concepts.—If we were to ask a highly specialized worker who was not familiar with or

concerned with the waves on either side of the ultra-high frequencies what electricity was like, he would be justified in answering as follows: Electricity is a phenomenon of waves of electromagnetic fields filling the space between conductors. The conductors and dielectrics serve to guide or store the energy. In transmission, either of two alternatives is common. It may be guided by a series of metal pipes inside of which both the source and the load are to be found. In this way the metal pipes ensure that the waves either go to the load to be absorbed or are reflected back to the source; the waves do not pass through the metal. Or the energy may be transmitted without benefit of continuous guiding boundaries, except as a starting mechanism, by creating around the source properly shaped boundaries that act to draw the waves from the source and release them into free space for further propagation. Note the emphasis that this microwave worker places upon the guiding of waves with boundaries.

Let us press him for more of his physical pictures. What of current and charge and voltage and impedance? Does he use these terms at the ultra-high frequencies? What meaning do they have for him? Current and charge, he tells us, he conceives of in two rather distinct roles, depending upon whether or not there are present in the problem electrons free from conductors. When speaking of microwave current due to an electron beam, his thoughts seem to dwell a great deal on the instantaneous distribution of electrons in space and their various velocities. When we speak of charges and current flow associated with conductors, he puts those down with finality as boundary effects and as of great importance. The conductors that bound a region containing microwaves have surface phenomena which he describes as currents and charges and which he appears to know if he knows the space distribution of the microwave fields and vice versa. He begrudgingly adds, in an aside, that the surface currents actually are not entirely on the surface; they penetrate somewhat, but he

does not seem to consider this phase worth elaboration. He makes a great deal of the idea of current densities and talks much of current densities in free space even when there are no moving charges there.

Voltage? A very important concept, he explains. It turns out that often the configuration of the region in which the microwaves flow is such that each point of the space does not have to be investigated or taken into account. Certain over-all or integrated effects, like voltage differences, are sufficient and often yield a superior physical picture.

At the mention of impedance, a gleam enters his eye; he seems to apply the impedance concept to everything. He

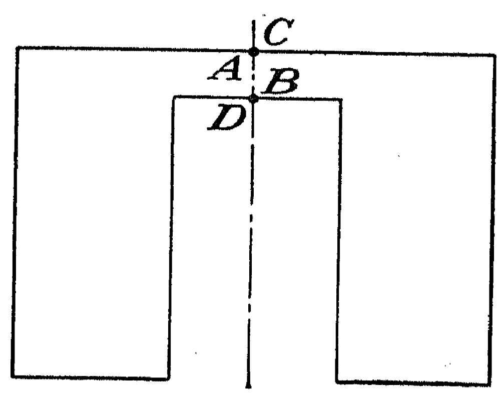

Fig. 3.

speaks of a point in space as having a certain impedance. A moment later he has changed the boundaries of the region slightly, and now he tells us that the same point has a different impedance. We ask, "What about the impedance of a two terminal circuit?" because that was what we really had in mind. He nods and points to two points A and B on the inner surface of a metal-enclosed cavity shown in cross section in Fig. 3 and says there is an impedance there. But a few mils away from A and B, on the outside of the shell, at C and D, there exists, he maintains, a radically different value of impedance. Apparently his idea of impedance is not exactly identical with that of a similarly fictitious worker who has known only low frequency.

We have looked first at microwaves as being the same as lower frequency electricity. This was done by observing that the ultra-high frequencies obey the same fundamental laws of electricity and magnetism. We have also looked at microwaves as being much different from lower frequency electricity by suggesting the relative inability to describe many microwave phenomena in terms of some narrow concepts so common at the lower frequencies. We draw from this the conclusion that, as we take up in successive chapters the characteristics of microwaves, it will be wise to examine each concept for its frequency breadth. It must be clear in the end that some concepts can be plainly set forth that will serve correctly over the whole region. We must understand how the pictures of the microwave and lower frequency engineers merge if both use broad concepts, and what approximations separate them if their notions disagree.

CHAPTER III

Microwaves Practically Do Not Penetrate Metal

Perhaps the first concept that should be clarified at any frequency is that of current flow in a conductor. We will first state what the situation is at the higher frequencies, then look into the explanation and implications of the result. Skin effect becomes so important as the frequency reaches the range of billions of cycles per second, the current comes so close to flowing entirely on the surface of conductors, that below a few thousandths of an inch there is hardly any current density left worth talking about. In other words, skin effect reaches its ultimate, and electric current in conductors becomes simply a surface phenomenon. Electric currents then do not flow *in* conductors in the microwave region; they flow *on* conductors for all practical purposes.

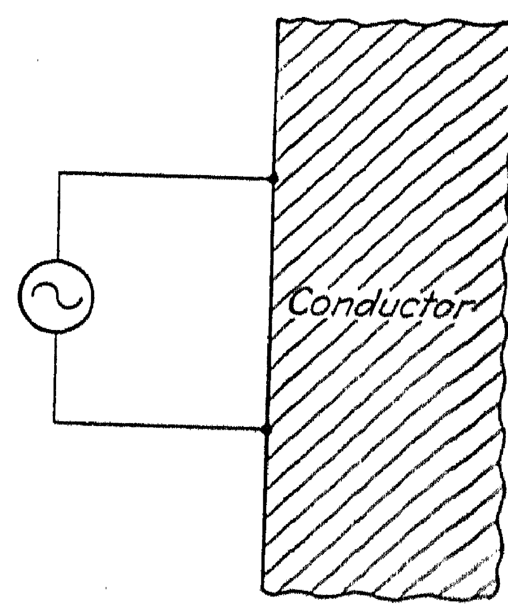

FIG. 4.—A conductor whose surface is part of a high-frequency circuit.

This is not the only phase of this surface tendency. It will shortly be seen that not only current flow but all manifestations of electricity and magnetism will fail to penetrate into the conductor. (The previous statement is completely true only for a truly perfect conductor, one of infinitely great conductivity. But in a practical sense, for copper or brass or silver, the extent of the actual penetration is so slight that it does not interfere with the points being made.) If somehow a good conductor, one whose surface is exposed to ultra-high-frequency excitation (Fig. 4), could be probed, then it would be discovered that electric fields

and magnetic fields as well as currents fall off extremely rapidly.

The implications of these results are broad, but first an examination of the reasons for these conditions is in order. For this, the fundamental principles of electricity and magnetism that will yield the correct answer at any frequency are used. Current flows at any point in a conductor because there is an electric field as the initiator of the current. Some readers will not be accustomed to thinking in terms of electric field as the cause of current flow at a point in a conductor, but rather of a voltage difference between *two* points. It is preferable at this stage to speak in terms of densities rather than over-all effects, *i.e.*, current densities rather than total current, and electric field at a point rather than the integrated electric field or voltage difference between two points. But knowing the impressed electric field at every point does not immediately yield the current strength there. In any problem, the electric field at any point is generally not simply the impressed field.

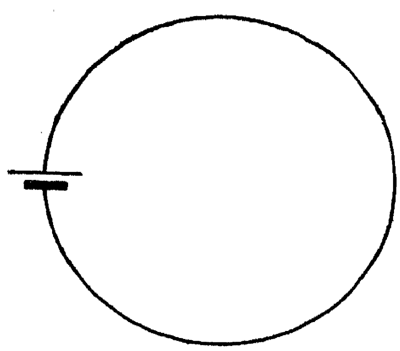

FIG. 5.—D.c. voltage applied to a uniform long conductor results in uniform electric field and uniform current density over the cross section.

Direct current is an exception. Here, when a voltage is applied to a system, the total electric field everywhere is the voltage drop per unit length, or the gradient. Apply the d.c. electric field uniformly, as in the long uniform conductor of Fig. 5, and the current flow will be uniform. But when the applied voltage and the resultant electric field vary with time, that field is no longer the total entire electric field, for the alternating current's magnetic field gives rise to an additional induced electric field. This effect, stemming from Faraday's law,[1] may be called an "internal inductance effect" in the conductor because it

[1] Faraday's law says that voltage is induced in a circuit if the magnetic flux linking the circuit varies with time.

uses up in a reactance drop some of the applied electric field, leaving only the balance for the ohmic resistance drop. We have now the basic principles: The current must distribute itself in a conductor so that the current density is given at every point by the conductivity times the net or total electric field. This total electric field will consist of one part applied by some source external to the conductor, and one part induced by the changing magnetic flux caused by the current.

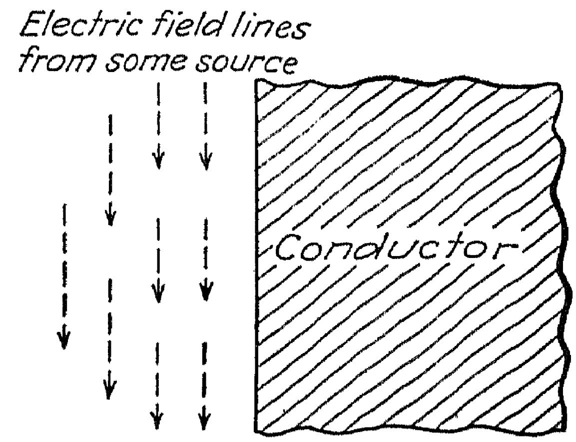

FIG. 6.

Perfect Conductor.—The process discussed above needs to be illustrated by some examples, and we choose, first of all, the case of current flow in a slab of perfect conductor. For our purposes a perfect conductor is an easy example; we will shortly add a small amount of resistance to it as befits common conductors such as brass or copper. Figure 6

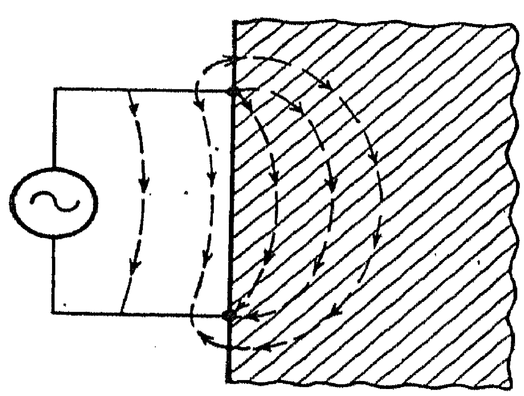

FIG. 7.—A source connected directly to the conducting slab tries to produce an electric field in the conductor.

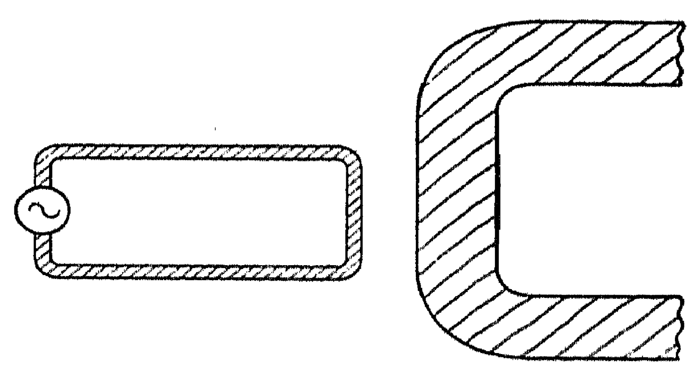

FIG. 8.—The source of electric field that produces current in the conductor may be an adjacent current-carrying loop.

shows such a slab and a rather general source represented by the lines of electric field. The source might actually be connected directly to the slab as in Fig. 7, it might be a loop as shown in Fig. 8 with the slab as part of another loop, or it might be the field of a distant antenna. In

all cases the idea pictured by Fig. 6 is applicable. The first thing noticed in the detailed analysis of this situation is that the total electric field tangent to the conductor's surface must be zero. If it were not zero, then there would be infinite current flow on its surface, an inescapable conclusion since the conductivity is assumed to be perfect, *i.e.*, infinitely great. Indeed, this must be true for the entire volume of conductor: the currents that flow in the conductor must cause just the right magnetic-flux distribution about them so that there will be induced electric field of exactly the right amount to buck out the applied field. A failure for the perfect conductor to defend itself in this way, an attempt on its part to have other than the correct current distribution, will surely result in a slight difference between the applied and the induced field at some point. That difference will immediately work on the zero resistivity to produce an infinite current. Such is known to be the wrong answer, an unnecessary answer certainly, that will cause infinite magnetic fields, that will in turn cause infinite induced electric fields, that will be even more physically impossible than the finite field. No, we reject the infinite current and hold to the first suggestion of zero total electric field.

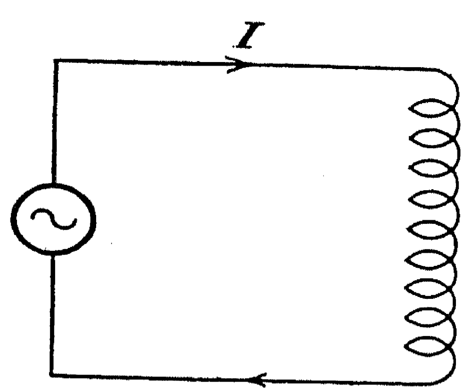

FIG. 9.—The current in a dissipationless coil is just enough to yield a reactance drop equal and opposite to the applied voltage.

The reader will recognize in this the same reasoning that he would apply to the case of a coil of pure inductance (all resistance assumed zero) driven by an alternating voltage (Fig. 9). The IR drop is zero since the R is zero. The current I is that value which will yield a reactance drop, really the induced voltage, which is just equal and opposite to the applied voltage. The total or resultant voltage is zero and is equal to the zero IR drop. What we did with the electric fields in the slab was precisely the same but on a differential scale. We examined points in the conductor,

instead of the integrated effects like the voltage across the reactance coil.

So much for the electric field for the moment. We see now that no electric field can penetrate a perfect conductor. What about the magnetic field? And, especially, what about the distribution of the current? That is the question. But the chapter is yet young.

Figure 10 shows a guess at the current distribution in the conducting slab and, likewise, the magnetic field linking it. We admit by this guess that we do not yet know how the current is to be distributed and consider a more or less uniform distribution as a starting point. Let us give our attention first to the magnetic flux. In particular, consider the amount of it that links the dotted rectangular loop *abcda* shown in Fig. 10. The rate of change of magnetic flux linking this loop will result, according to one basic law of electricity and magnetism, in an induced electromotive force around the loop. In other words, granted that there is no electric field on the surface, this electric field would have to increase with depth into the conductor because the changing magnetic flux is contained between the surface and the interior. For example, in the loop shown, since the electric field on the surface along *ab* is zero, there would have to be electric field along part or all of the remainder of the loop *bcda*. Moreover, there would have to be enough electric field so that when it is summed up over *bcda*, it will add up to precisely the induced e.m.f. that the changing magnetic flux is causing.

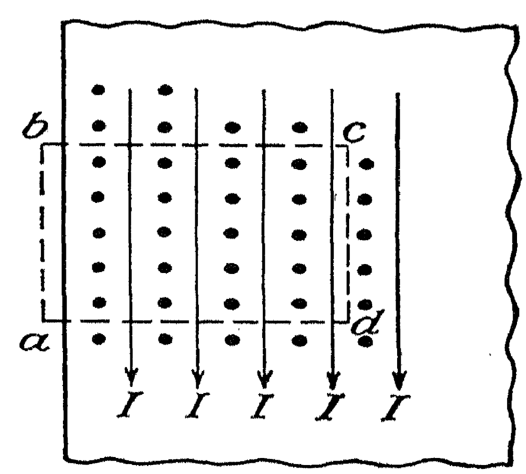

FIG. 10.—Cross section of part of the conductor. The arrows represent current flow parallel to the plane of the section. The dots represent magnetic flux lines caused by the current, these lines being perpendicular to the plane.

But what we have described leads us to an impossible condition, *viz.*, electric field inside a perfect conductor—the situation that, as before, would lead to the absurd

infinite current. Thus we conclude that the magnetic flux cannot penetrate a perfect conductor. As soon as we assume, as we did, that magnetic fields can exist inside a perfect conductor, we are led by reasoning from correct fundamental laws to the impossible infinite currents. It is worth recalling that the magnetic fields we are speaking of here are those which are varying with time. A d.c. or, better, a static magnetic field will penetrate a good conductor, for, without the induced electric field that comes from the time change in the flux, there is no means for the conductor to resist the magnetic field.

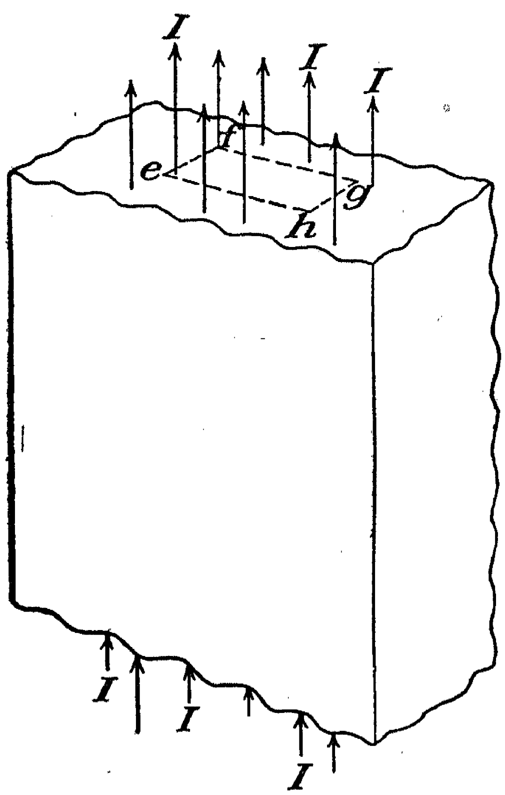

FIG. 11.—If there were current flow (indicated by the arrows) inside the conductor, magnetic field would have to exist around all or part of the loop *efghe* which links the current.

Now we can observe almost at once that no currents that vary with time can penetrate a perfect conductor. If the currents were present throughout the volume, then (again, according to a fundamental law that holds equally well from direct current through microwaves) there would have to be magnetic fields accompanying them. It might be argued at first that, though the currents are each encircled by magnetic-flux lines, perhaps the correct current distribution might work out to be such that the separate magnetic fields of each current filament oppose one another and give a total or resultant magnetic field of zero everywhere in the conductor. Actually, no such current distribution could be found. This can be seen by a proper application of the fundamental law. Figure 11 shows again the imagined current distribuiton in the slab, and we suppose that the magnetic field is really zero, as required, at some part of the interior, say along *ef*.

But if there is current in the interior of the conductor, the loop *efghe* will encircle some current. The law relating current to magnetic field tells us that some magnetic field must exist somewhere or everywhere along the remainder of the loop *fghe*. There must be enough magnetic field around this loop so that, when summed up around the loop, it will yield a magnetomotive force (m.m.f.) that is proportional to the current encircled. This is simply a statement of the basic law, and there is no arguing with it. Thus, for the changing magnetic field to be zero everywhere in a perfect conductor, the current density must also be zero everywhere.

We are left then simply with this picture of a perfect conductor: When a perfect conductor is subjected to applied electric fields or changing magnetic fields, enough current flows on the surface to produce electric and magnetic forces of its own that are just sufficient to buck out completely all effects inside the conductor itself.

Imperfect Conductor.—The slightly imperfect conductor will tolerate a small net electric field on the surface instead of insisting on an absolute zero value there, as the perfect conductor did. Again, this is all determined by a condition of balance predictable by the fundamental laws; there will be enough electric field left to overcome the resistivity of the conductor and no more. Figure 12 shows a plane conductor of great depth subjected to electric field on its surface. The total electric field does not have to be zero anywhere; thus a certain distribution, as yet undetermined, is allowed throughout the depth of the conductor. But the total field, and consequently the current, will not be uniform. If it were, then, according to the argument used previously, the changing magnetic flux would yield a continued rise in induced electric field with depth that, after overcoming the applied field, would nevertheless keep increasing. Then the current would increase indefinitely the farther away it got from the source. This is the result of guessing at a uniform current distribution. What

is more reasonable physically, and in complete agreement with the basic laws, is that an imperfect conductor will (1) allow enough current penetration (2) to yield enough magnetic-flux penetration (3) to cause just enough induced electric field inside the conductor (4) to overcome gradually the applied field. As the total field falls off, the current falls off, its magnetic flux falls off, and the change in induced field no longer is appreciable. It approaches the applied field in magnitude but is opposite in sign, the difference being the total field that approaches zero with increasing depth.

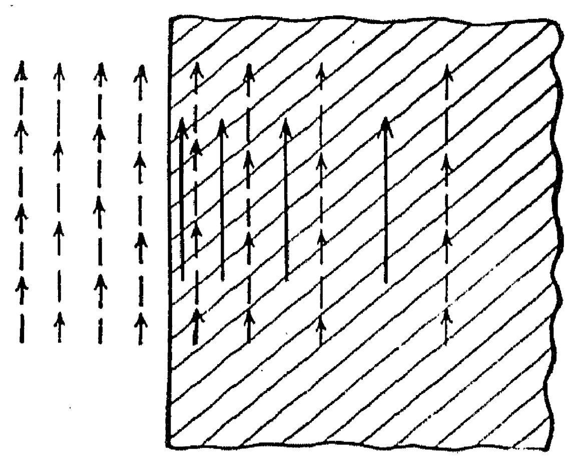

FIG. 12.—In an imperfect conductor, the electric field (dotted arrows) does not have to be zero inside the conductor, but both the electric field and the current density (solid arrows) fall off rapidly with penetration into the conductor at very high frequencies.

Now, how quickly do the currents and fields fall off in strength as the surface is penetrated? For any given conductor, this depends upon frequency. It is the rate of time change of magnetic flux that is responsible for the marked change in current distribution with frequency. A given current filament will produce almost the same strength of magnetic field in its vicinity over the whole frequency range. Therefore, it is not this factor that brings in the frequency effect. But if the frequency is high enough, the magnetic-field change will be easily able to produce an induced field high enough to buck out the applied field inside the conductor, and the only current flow will be near the surface. If the frequency is low

enough, the effect is so slight that the change in current distribution from that of direct current may be scarcely noticeable.

Thus the main findings about skin effect may be summed up:

At low frequency, the current distribution in everyday conductors is uniform, with modifications to account for the fact that the current is alternating.

At the ultra-high frequencies, the current is entirely on the surface, with modifications (a certain depth is allowed) to account for the finite resistivity. Furthermore, fields fail to penetrate a good conductor for all practical purposes. And, finally, the surface under consideration is not necessarily the outside surface of a conductor but simply that surface exposed to the source.

These conditions have a very important effect upon the engineer's thinking. Almost the first facts we want to know about centimeter waves are where they are, where the energy is being stored, and toward where the power is being guided. If practically nothing goes on inside conductors, if only conductor surfaces take part in the action, then the conductors become simply boundaries, enclosing regions in which the centimeter waves really do their propagating and existing. Accordingly, it will readily be understood why the microwave engineer in his thinking uses pictures of electric and magnetic fields a good deal of the time instead of working entirely with current and charges or voltages of the system.

This does not mean at all that the concepts of current and charge and voltage have become obsolete. Rather, these concepts share the limelight with fields in situations where previously the field concepts have been more in the background. That sharing often brings a broader instead of a narrower appreciation of the current and voltage concepts.

Of course, at any frequency, when circuit impedances, inductances, and capacitances are dealt with, somewhere

back in the analysis there must have been a consideration of the electric- and magnetic-field distribution; *i.e.*, the familiar picture of the relations between current, voltage, and charges of electrical systems ultimately depends upon linkages between those systems that come about because of their fields. In the usual low-frequency physical picture, the electric and magnetic fields are feelers that emanate from the centers of electromagnetic effects and are useful in accounting for the interaction between different conducting systems. But, in the physical picture of the engineer, the center of the problem at the lower frequencies is generally in the conductors and the current flow in those conductors. Two common examples in the microwave

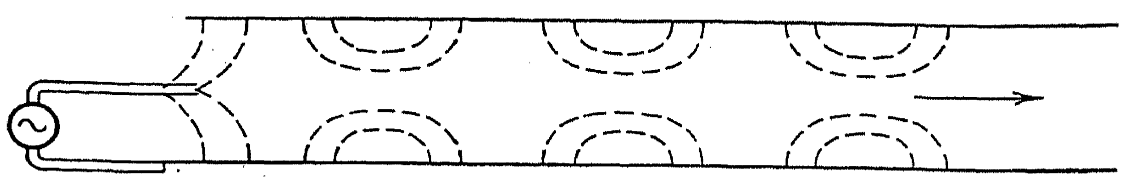

FIG. 13.—If the frequency is high enough, currents, charges, and electromagnetic fields can be made to propagate down the inside of a hollow conducting cylinder.

frequency bands illustrate how the center of the problem expands from the conductors' voltages and currents and charges to the region surrounded by these charges and currents in which electric and magnetic fields exist.

In Fig. 13 is shown a system that might be used for transmitting electromagnetic energy in the microwave region—a long cyclindrical conducting tube with a source of centimeter waves near one end. A small probe is charged by the high-frequency generator so that lines of electric flux from that probe end on the cylinder's inner walls. If the frequency is high enough, electromagnetic fields will propagate down the inside of the tube. This example will be expanded in a later chapter, but for the time being it is noted that what goes on in the tube may, of course, be studied with the spotlight on the current flow and the charges on the inner surface of the cylinder. Electric

currents flow, however, only on that boundary; no electromagnetic effects are able to penetrate the cylinder to the outside. Thus it is understandable that the situation pictured may be looked upon as one in which the setting of the problem is no longer in the conductor, but rather in the space surrounded by the conductor. The cylinder is looked upon as a wave guide. Figure 14 shows a hollow conducting box which is in some ways typical of resonant circuits at ultra-high frequencies. Again, this example will be studied more carefully under microwave circuits; but at this time it is assumed that there is a source of centimeter-wave energy inside the box and that as a result there is current flow on the inside surface. The top and bottom of the box become oppositely charged, and currents connecting these charged top and bottom plates flow up and down the walls of the box between those plates. Again, since no effect exists outside the box, because of the negligible penetration of current at these frequencies, and since the box walls are boundaries of the problem, the focal point of the problem appears to a great extent to be the space enclosed by the box rather than the conductor itself.

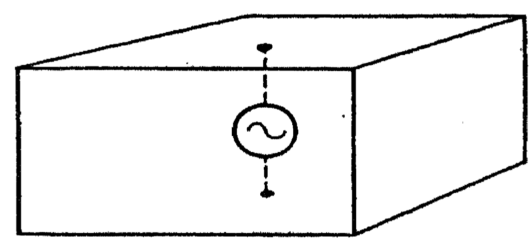

FIG. 14.—At high enough frequencies, currents, charges, and fields may be made to exist on the inside of a cavity bounded by a conductor.

In a much broader sense than that to which engineers who have not worked with centimeter waves will have become accustomed, current flow becomes a boundary condition—not the central core in the picture. This begins to make electric effects in the centimeter-wave region allied closely with light waves, which are usually considered as passing through a medium with the boundaries acting as absorbers or reflectors but, nevertheless, boundaries only. Light is ordinarily thought of as a phenomenon that takes place in a medium that will *transmit* it—not as something that takes place *along* mirrors or *along* other boundary surfaces.

This picture of centimeter-wave current flow along conductors will be discussed later when circuits and transmission lines are considered. For the present it will be sufficient to give free rein to the idea that at these extremely high frequencies there is practically no such thing as current flow in conductors. The charges and currents of a conductor sit and flow on the conductor surfaces.

CHAPTER IV

Electrons Travel Slowly; Thus Transit Time Becomes an Important Concept

It has been stated that this text is not concerned with anything but basic fundamentals and physical pictures. Yet the workings of an electronic tube are to be discussed next. The reader is justified in asking why what takes place in an electronic tube, presumably a "device," should be introduced here as a fundamental concept. The answer brings out certain important characteristics of microwave electricity.

At the lower frequencies the vacuum tube is perhaps well described as a device. It is something which can be inserted into a circuit and made to perform many unusual jobs not possible with simple circuit elements such as R, L, and C. The electronic tube is, however, the beginning and the end to the centimeter-wave engineer, and tubes are so closely tied in with circuits or wave guides that any concept of current flow is poor unless it correctly pictures microwave current flow from conductors into electron streams and back to conductors again. A large portion of the time, the current of the system is in raw electron streams in vacuum rather than in current flow along conductors. So it can be said that it is not as a device that the electronic tube is being discussed; rather, what is being considered is the other basic type of current flow: microwave currents in a stream of free electrons. The question, however, is even broader than that.

At low frequencies, or power frequencies, to be exact, two very fundamental principles are those dealing with (1) the production of mechanical forces on current-carrying con-

ductors and (2) the generation of voltage by mechanical motion. These two principles are behind the motor and generator, which are two important, if not the very most important, devices of the power-frequency range. In the microwave field, it can probably be said again that almost everything starts or ends with transfer from or to mechanical forces. Moving electrons, however, and not rotating machines, are the object of these forces. By proper utilization of the mechanical energy in the moving electrons, the high frequency is generated. By properly affecting the motion of electrons with forces produced by electric or magnetic fields, the high frequency is utilized to accomplish the purpose for which it was generated and transmitted.

In this input and output process in which electrons are concerned, it is not always so important to be concerned with mechanical forces and the conversion of motion into currents. For instance, at ordinary radio frequencies the over-all effect is sufficient; curves are available giving currents passed by the vacuum tubes in terms of voltages applied to the electrodes. More than this, the explanation of the operation of the tube deals almost exclusively with these over-all voltages and currents and does so successfully. But, unfortunately, as the frequency creeps up into the hundreds and thousands of megacycles, the pleasure of avoiding a look at the fundamentals of electron streams becomes a risky one in which to indulge. The two phases of the phenomenon must be investigated. This chapter will take up the influence of ultra-high frequency on electron motion; the next, the way in which electron motion gives rise to electric currents in adjacent conductors.

Review of Fundamentals.—In our first studies of electricity, after we mastered the cat's fur and amber, the hero of the next chapter was the electric charge. Like charges repelled, and unlike ones attracted each other, we were told; and before long the idea of potential difference between two points was introduced as the work done in moving a

unit charge between the two points. Work had to be done because of the force felt by that unit charge due to the other charges. In other words, the unit charge was said to be in the "field" of the other charges, whose field strength was defined as the force on the unit charge. In this way our first pictures of electric field and potential difference or voltage were built up.

This is all summarized in Fig. 15. Here are shown two plates across which there is a steady difference of potential. Because of this d.c. voltage there are positive and negative charges sitting on the two plates, and the charges are thought of as connected by electric-flux lines. There is an electric field between the plates, so that if a unit electric charge, like Q, is placed there, a force will be exerted on it. If the charge is positive, and if it is moved forcibly from the negative to the positive electrode, a certain amount of energy is necessary for the task—an amount equal to the potential difference between plates, that being after all the definition of the potential difference. The work done can be calculated also, by figuring that the charge is moved against the forces of repulsion from all the positive charges on one plate, and against the attraction of the negative charges on the other. This is one way of stating the criterion that dictates how much charge should be on both plates: just enough everywhere so that the work done in moving a unit charge between them, no matter what path is chosen, will equal the applied voltage.

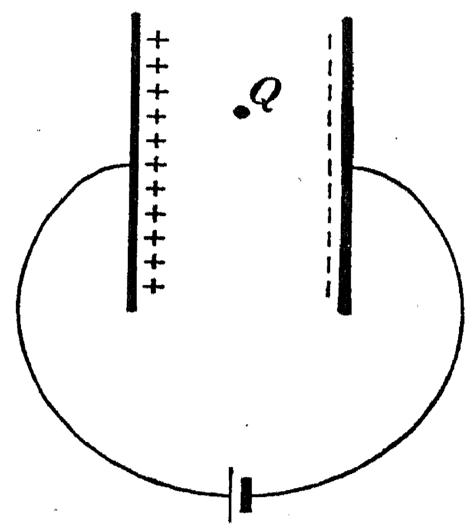

FIG. 15.—The work in moving a unit electric charge between plates is equal to the voltage difference between them.

If a negative unit charge is released at the negative plate, then here obviously no work will have to be done to move it to the positive plate. The attracting and repelling forces of the charges on the plates will now be in the direction of motion. If the charge starts from rest, it will be accelerated

by the field; *i.e.*, work will be done on it. At each point, its kinetic energy of motion is equal to the difference of potential between that point and the starting point. If it had an initial velocity, then the statement is simply revised: The *increase* in kinetic energy of motion is equal to the described difference of potential. Thus is acquired the concept of an electron's picking up a voltage when it passes between two electrodes having that voltage difference. Applying voltage to a tube's electrodes is a way of imparting velocity and energy changes to electrons that cross the spaces between them. In fact, it is common to give the increase in electric energy in terms of electron volts and to name only the accelerating voltage responsible for the increase.

This review of fundamentals is enough to develop satisfactory physical pictures of the operation of essentially all electronic tubes at low or medium high frequencies. Thus, in the simple triode, the usual explanation is about as follows: The cathode emits an abundance of electrons, and these would normally flow hurriedly to the plate that is maintained at some higher positive voltage with respect to the cathode, were it not for the effect of the space charge of the electrons themselves, aided by the presence of the negatively biased grid; *i.e.*, the electrons when released fill the space in the region between the cathode and the grid. These emitted electrons, together with the grid, exercise a repelling force on the electrons that have just been emitted and are about to start toward the plate. As a result only a certain amount of electrons, those emitted with sufficient velocity, will be able to migrate to the plate. The first part of the picture, then, is that current flows from cathode to plate; the amount, however, is not limited by the ability of the cathode to emit—of course, that is a possibility, of little importance now—but rather by a retarding force in the path of the electrons.

The more the plate is made positive in voltage compared with the cathode, the greater the number of electrons that

will have energy enough to pass the critical region and travel to the plate. The grid is located closer to the heart of the bottleneck, and thus the velocities of the electrons are very sensitive to its voltage. The usefulness of the triode follows from these properties. The tube manufacturer presents a neat set of curves giving the relation between the various currents and voltages. To design a common resistance-loaded amplifier, as shown in Fig. 16, an inspection of these curves would be made. From them would be obtained all the necessary data to calculate the amplifying ability and power output of the circuit.

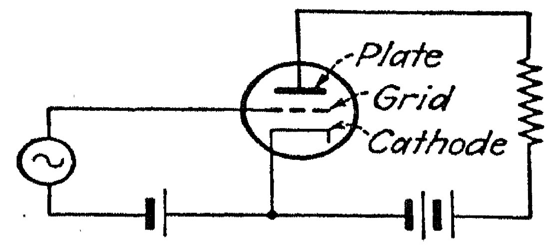

FIG. 16.—A change of grid-cathode voltage causes a change of current in the output resistor.

Electron Transit Time.—Throughout the preceding explanation, no mention was made of the time it takes before a change in grid-cathode voltage can exert its effect on the space-charge region, or of the time it takes electrons to dash from one position to another and pick up velocity in compliance with the orders of the applied signal. Instead, it was assumed in the foregoing description, just as it is assumed most of the time from the power frequencies to well beyond the broadcast band, that for every set of values of plate voltage and grid voltage there exists a certain value of plate current. Moreover, it was taken for granted that if the grid voltage or plate voltage or both are changed to new values, the current will immediately change to a new value, without regard for the speed with which the changes in voltage may have been made.

If a rapidly oscillating voltage is added to the d.c. biasing voltage on the grid (Fig. 16), then it is known that the space charge cannot jump instantaneously to an oscillating condition in exact step with the grid voltage. Nor will the plate current be caused to vary periodically in synchronism with grid voltage in such a way that the instantaneous value of plate current will be related to the

instantaneous value of grid voltage as though both were static. It is known that the application of the oscillating voltage on the grid will result in a periodic attraction and repulsion of the electrons toward the grid. That much is true regardless of the frequency. The result will be that the space-charge electrons near the cathode will commence to move under the influence of this newly applied force. But the existence of the force does not guarantee a change in velocity. According to the laws of mechanics, the force, as soon as it is applied, causes an acceleration. This acceleration results in more or less velocity, after a period of time has passed, depending upon the mass or inertia of the electron. It takes time, in other words, for the electron to accumulate the energy increase which the voltage holds out to it. If the voltage would only stay at the same value until the electron has completed its travel through the whole region of influence (as it does for direct current, for example, and essentially, also, for low- or medium-frequency alternating current), then and only then would it pick up the energy at each point that the voltage difference dictates. The reasoning from electrostatic theory of a few pages back would then hold. However, at ultra-high frequency that reasoning does not hold very well.

Ordinary electrons travel much too slowly compared with the period of the alternating current. Consider most triodes and an operating frequency in the billions of cycles per second: When some emitted electrons, encouraged by the grid going more positive with respect to the cathode during part of the cycle, have received a little more energy and have gone a small part of the way toward driving through the space-charge barrier, the grid voltage may reverse, slowing the electrons down again and perhaps actually forcing them to return to the pile of surplus electrons in front of the cathode.

It cannot be assumed that this effect of electron transit time is nonexistent. To be sure, this does not mean that

tubes stop working and that ability to control electrons simply fades away at ultra-high frequency. But it does mean that it is necessary to keep *transit time* in mind. Perhaps new tubes are invented that make use of transit time as a very basis of operation, or higher average velocities are used, or spacings between electrodes are decreased; but it is a general rule in any tube useful at extremely high frequencies that transit time is a powerful factor in deciding the operation of the tube. The fundamental reason is that about all to be hoped for with an electron is to change its velocity either in direction or in magnitude. If a tube is designed with transit time of electrons ignored, it may well turn out that the input energy is not noticed in any practical way by the stream of electrons and the performance of the tube deteriorates with increasing frequency. With some appreciation of the fact that an electron stream consists of relatively slow moving charged particles, ways may come to mind to make the input voltage effective again in influencing the electron stream and hence in producing some useful action in the tube's output.

Input Effect.—As for the main objective of accent on fundamental concepts, the issue here is the relation between cause and effect. The effect of a voltage applied to an electron beam is to change the energy of the electrons. This is an input effect. Presumably, whatever tube it may be, the rest of the workings of the device depends upon the electrons' first receiving that energy change. The magnitude of the effect for the same cause, *i.e.*, the same magnitude of voltage, is something that varies greatly with frequency. It is possible to use certain concepts derived from electrostatics for most of the useful frequency range to give the relation between applied voltage and electron-energy changes. These concepts are good while the frequency and physical dimensions are such that the electrons travel between electrodes quickly. Their transit times should be short compared with the period of the alternating current.

We continue this lesson with two illustrations. One is a striking but physically improbable case, discussed here to define a principle. The other is a practical case typical of a tube intended especially for centimeter-wave work in which transit time is used as a beneficial, necessary phenomenon.

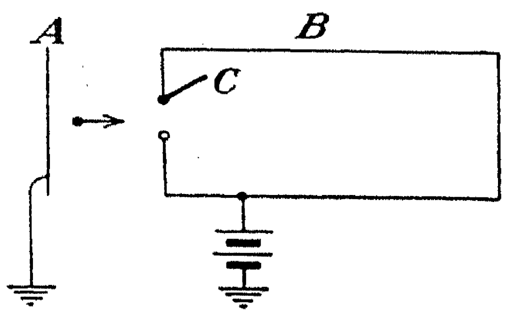

FIG. 17.

Figure 17 illustrates the first case. An electron is emitted at zero velocity by the cathode A and is accelerated by the d.c. voltage toward electrode B. At every instant, at any point of the path, the velocity of the electron can be determined—by making use of the proportionality between the kinetic energy of the electron and the potential of that point in space.

Suppose that the electron has traveled long enough to enter electrode B. Then, close the gate C, and electrode B is now a perfect shield. No electric field can be placed inside it from the outside, and thus no force can be exerted upon the electron traveling inside it. Figure 18 pictures the electron traveling inside B, but with the d.c. voltage source exchanged for a new alternating generator. Let us imagine this generator to be an extremely active and enthusiastic one. Let it pull the potential of B up and down with no reservations, over even greater voltage swings than the original d.c. voltage of B. Still, since no electric field acts on the electron, it will remain coolly intent on traveling with constant velocity throughout the length of electrode B, completely disregarding the electrical havoc taking place outside its shield.

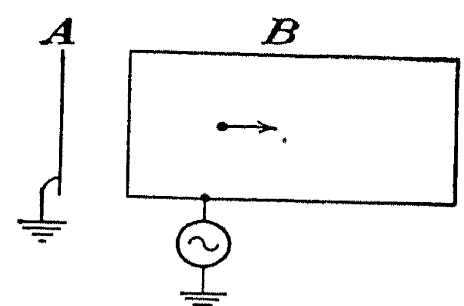

FIG. 18.—The electron inside the shield B does not notice the effect of the varying voltage between B and A.

When the electron reaches the end of B, with the same velocity it had when it entered, open the gate D (Fig. 19) and allow it to start experiencing such electric field as

may then exist between electrodes B and E. Just to steady things a bit, assume that the electron emerges at an instant when B is at ground potential and E is negative with respect to ground and that these two voltages will remain that way. Notice, then, that the electron is leaving an electrode at potential zero, if potential is reckoned from the starting point of that electron, and yet it has a definite velocity that bears no relation to the present potential of B. The field between B and E acts now to decelerate the electron, since E is negative compared with B and the electron is negatively charged. Nevertheless, if E is not sufficiently negative in potential, the electron may run right into it; *i.e.*, it may be collected by the negatively biased electrode E.

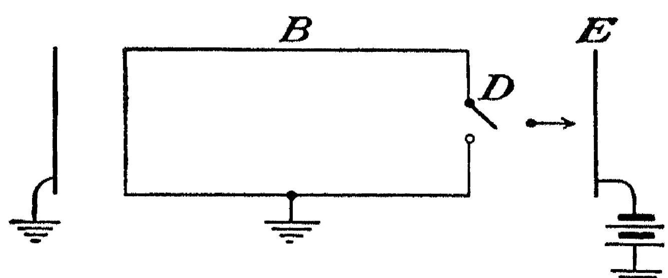

FIG. 19.—Electrons may be collected by negatively biased electrodes.

The situation discussed does not occur in any practical tube (at least, not in one complete with gates), but it magnifies a number of fundamental points in influencing an electron stream: The velocity of an electron is not set by the voltage established in the space where the electron finds itself. Its velocity depends on its whole background in time, starting from the instant the electron first began to move; and, relatively speaking, the electron may be traveling very slowly. An electron may be collected by a negatively biased electrode, *i.e.*, one negative compared with the source of the electron. An electron is affected, accelerated by fields; that part is true as a broad principle over the whole frequency range.

Velocity Modulation Tube.—The final example is that of the input end of a tube designed to use transit time as the basis of operation. The ideas behind this tube

may be called "ultra-high-frequency concepts" all by themselves. They would be impractical at the lower frequencies. They might even be said to be inconceivable at lower frequencies, if thinking about tubes is restricted to over-all voltage and current effects and if electron velocities and oscillations are overlooked. This tube is most often called a "velocity-modulation" tube. In one version of it (the input end of which is shown in Fig. 20), the electron stream is affected by the incoming energy quite differently from the triode that has been discussed. Here the electron beam is first accelerated to a high velocity and is contained mainly in a closed conducting structure that shields it from electric fields. However, at two

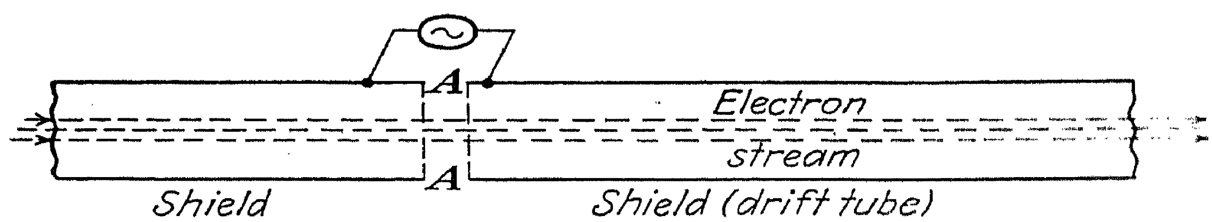

FIG. 20.—An electron stream can be velocity-modulated by applying alternating voltage across a gap in the stream's shield.

places, at least, along the length of the beam—one for input, the other for output—the shielding is broken.

At the gap in the shielding *AA* it becomes possible to introduce an electric field, and the electrons of the stream feel this electric field as they pass the gap. Imagine then that the input voltage is applied across the gap *AA* and that, being alternating, the voltage accelerates some electrons and decelerates others as they pass the gap, depending, of course, on the direction of the oscillating field when those particular electrons passed. If the frequency is high enough so that the electrons travel slowly, relatively speaking, they will be able to contemplate the influence they have felt as they move away from the gap down the drift tube (Fig. 21) and will have time to plan some useful output effects for the tube. Fast electrons, those which received accelerations when they passed the gap, will start to catch up with slow ones from the preceding cycle

(Fig. 21). This will cause the beam to develop dense regions, where the fast and slow electrons are tending to group, and rarefied regions, which these fast and slow electrons have deserted.

In other words, the velocity modulation of electrons at one point in their path AA causes an originally uniform density or d.c. beam to form into a modulated stream of electrons that at any later point contains a.c. as well as d.c. components. A suitable output device, which is left as an example for the next chapter, can abstract the a.c. current as a useful output. But notice how important it is that the frequency be high—or that the electrons move slowly. If the electrons travel the whole length of the

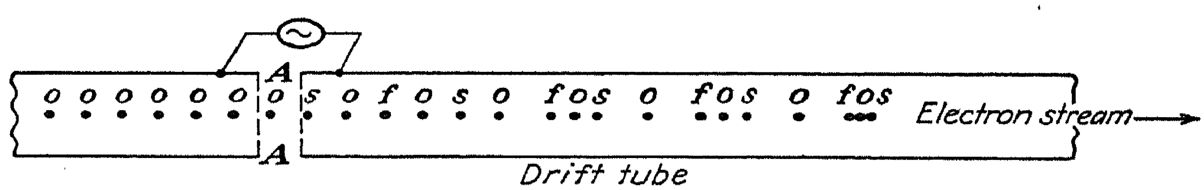

FIG. 21.—Faster-than-average electrons f (those that passed the gap when the input voltage was accelerating) overtake average electrons o and slower-than-average electrons s.

tube before the gap voltage has changed much, surely the faster electrons cannot hope to catch up with the slow ones of the preceding cycle. Those slow ones will have passed down the tube and been collected ages before, to speak in terms of the life of an electron in the space between source and collector.

Also, it is worth noticing that in the input gap itself the question of transit time enters in an important way. If the gap is too long, the transit time too great, with many reversals of field in the space before an electron passes by, that electron will have to add up the contesting accelerations and decelerations to a resultant that may be very small. Again, even in this transit-time vacuum tube, the input voltage is not implanted completely in the electrons.

Summary.—In this chapter it was recognized, first, that the motion of electrons, as affected by ultra-high-frequency electricity, is important. Then it was pointed out that

it is not sufficient to think of an electron tube as a device that passes more or less current as a function of the various instantaneous electrode voltages. The transit time of electrons enters at the higher frequencies to such an extent that the ability to influence electron motion with applied signal voltages between the electrodes may deteriorate rapidly. Certain ideas, such as the one that an electron's velocity depends on the potential of the space where it happens to be, were criticized as often completely inapplicable to high frequency.

It was necessary to return to fundamentals, forces on charges due to electric fields, and consideration of all the accelerating forces during an electron's flight to determine the net effect of an attempt to influence its motion with applied voltages. The problem of finding precisely what happens when high frequency is applied to an electronic tube became complex—a combination electrical-and-mechanical problem. Yet it was possible to see that, because electrons generally travel relatively slowly in the microwave bands, the operation of simple tubes may change materially. Also, it became apparent that it is possible to design new tubes that actually utilize transit time as a basis of operation.

CHAPTER V

Moving Electrons Induce Current

When electric current is made up of a stream of electrons between electrodes of a tube, it would appear to be the simplest, most straightforward current flow possible. In a conductor, current flow is rather complex at any frequency. The mechanism itself, *i.e.*, whether the current is due to a passage of electrons from one end of the conductor to the other, how the electrons move about in the atoms of the conductor, etc., is difficult to discuss without bringing in all the complex tools of modern mathematical physics. But current caused by an electron stream must surely hold the answer to a student's prayer. The flow is merely a motion of electric charges; it is the very definition of current in action. If one knows the density and velocity of the electrons, the current is known. How can there be any more to it than that?

But the matter is more complicated. Thus a chapter is devoted to currents due to the flow of electrons. First of all, there is a concept that must be discussed in detail, one that is excellent for d.c. and the ordinary power and radio frequencies and is accordingly in common use there. However, for ultra-high frequencies that concept may be too inaccurate to be practical, and it will often give entirely incorrect answers. When this is shown to be true, then the inclusion of this chapter will have been justified. In the previous chapter, the first part of the electronic-tube problem was discussed, *viz.*, the influence on the electrons of the applied voltages or electric fields. Now it must be assumed that the electron stream has been modulated by one means or another, with the problem remaining to

find the currents in the surrounding conducting structure that result from the motion of electrons.

The discussion can perhaps best be started by considering a single electron passing between two conducting planes. In Fig. 22 the electron is first near the electrode at the left, with lines of electric flux from it falling upon both of the electrodes and ending there on positive charges that will be induced on the plates to match the negative charge of the electron. The ratio in which the positive charge

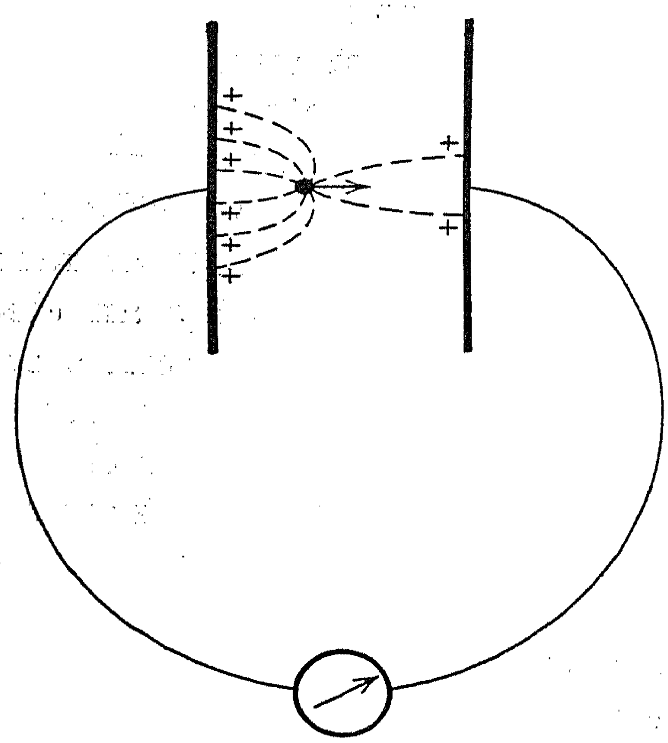

FIG. 22.—As the electron moves, the positive charge is transferred from one plate to the other.

divides itself between the two electrodes is dependent, of course, upon the adjacency of the electron to these electrodes. At first, with the electron very near the electrode on the left, most of the induced positive charge is on that electrode; but, as the electron moves, a greater percentage of that positive charge appears on the plate to the right. This means, of course, that in effect some of the positive charges from the left electrode must be moving through the ammeter, appearing on the electrode at the right at just the correct time to receive the flux lines that have shifted over. As the electron moves between the

two plates, the amount of charge on the left electrode decreases, and the amount of positive charge on the electrode at the right increases. Consequently, throughout the motion, a current, which is a measure of the rate of change of charge, flows through the ammeter.

Finally, as the electron nears the electrode at the right, all the positive charge, or essentially all of it, now accumulates on that electrode. At the very end of the trip, when the electron finally lands at the right plate, it meets the equal positive charge resting entirely on that plate, with no charge whatsoever remaining on the opposite plate. The collected electron neutralizes that positive charge, and the flow through the ammeter becomes zero.

Current flows all the time during the electron motion. Indeed, it stops only when the electron has finally arrived. From this picture it is seen immediately that it is not necessary for an electrode to be collecting electrons in order for current to flow from that electrode to any other. It must receive electrons if it is to pass a direct current, but instantaneous current will flow while the electrons are merely approaching the plate. Had the electron turned around and gone back before it had completed all of its travel, then the instantaneous current would have reversed, but it would have been zero only at those times during the process when the electron velocity happened to be zero.

A safe and correct viewpoint is always to consider the current in an electronic tube as arising from induction because of the motion of charges in the space between electrodes. The time when the electrons are finally collected is not the time when current flows; as a matter of fact, that is the time when current due to those electrons ceases to flow.

If there is a continuous stream of electrons coming from a cathode, say, to an anode, then, as each electron moves across, it transfers the matching electric charge present on the face of the electrodes from the cathode to the anode through the external circuit. A continuous process, a

steady flow of electrons, means that the current flowing in the circuit is equal to the number of electrons arriving at the plate per second. In this case, the two concepts agree. Current flow in a vacuum tube may be regarded as due to the instantaneous collection density of electrons to the electrode, or it may be regarded as a movement through the external circuit of the induced charges on the electrode, occasioned by the electrons' presence and motion. For either concept the answer is the same. But notice that, in the absence of this regularity, the number of electrons leaving the cathode at any instant would not necessarily equal the number of electrons arriving at the plate at that same instant. Particularly does this statement have significance if the frequency is so high that the time taken by electrons to move across compares with the time for a complete electrical cycle.

Current Flow Due to Motion of a Distribution of Charges. The centimeter-wave engineer, when he thinks of current flow through an electronic tube, cannot overlook the fact that the current flow is due to the motion of a distribution of charges, and it is an integrated effect of all the induced current due to all the charges with which he must deal. A few examples from practical tubes will illustrate this.

First, consider the triode again, and assume that the grid is negatively biased. That, it is understood, means that none or few electrons will be collected by that electrode. No direct current flows to the grid from the cathode simply because no charge is collected. Whether it can be said, in addition, that alternating current flows from the cathode to grid requires some study. Surely at a low frequency, if not at very high ones, this lack of collected electrons would be sufficient reason to conclude that there is no a.c. current flow into the external circuit. But, in view of the discussion just completed, the fact that the grid is immersed in a region where there are moving electrons cannot be overlooked. The broad concept must be used to make sure whether there is or is not a.c. current flow. So it

will be necessary, in principle, at least, to add up the separate inducing effects of all electrons that are moving about within electric-field-reaching distance of the grid.

All this sounds as though the number of electrons and, indeed, the position and velocity of each electron must be known. A simplified example will not interfere with a demonstration of principles, however. Figure 23 shows a triode with planar electrodes of large area compared with the spacings between them. The grid is assumed to be finely spun so that the space-charge action, a.c. or d.c., is about the same no matter which portion of the cathode is considered. Suppose now that the shading in Fig. 23

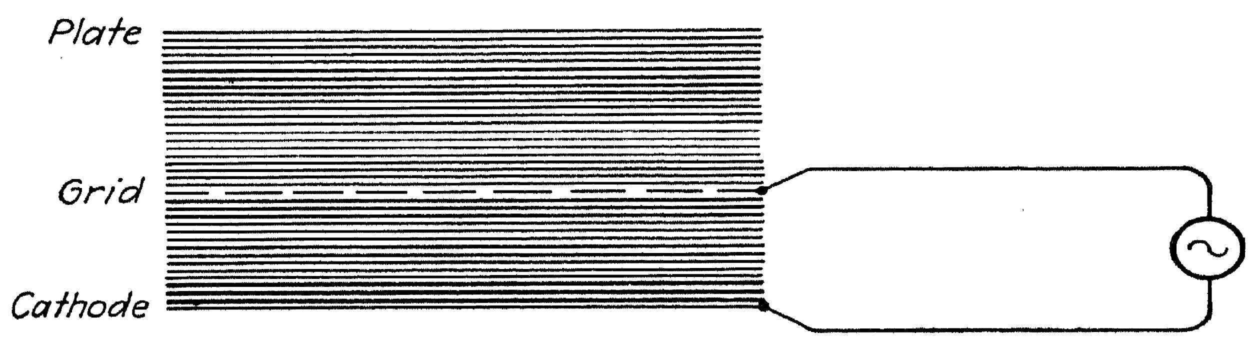

FIG. 23.—Very high frequency voltage applied between grid and cathode results in density waves in the space charge.

denotes an increase in electron density over the d.c. value. The variations in density are the result of the a.c. signal voltage that is applied between grid and cathode. The frequency is assumed to be so high that the voltage reverses itself before a space-charge increase (that was a result of the voltage increase) has had an opportunity to manifest itself as a change in plate current. Thus is shown a series of waves of space-charge density, all of which is superimposed on the direct current between the electrodes.

Notice that, though these waves have been drawn with no attempt at precision, certain effects have been carefully attended to. For instance, the space-charge variations are grouped together closely in the cathode-grid region, less closely in the grid-plate region. This is because the cathode-grid region is a relatively low d.c. potential region, while the plate-grid region is relatively high in d.c. potential.

In the grid-cathode region, the average electron velocities are lower and the wave effect, evidence of low transit time, is more pronounced than in the grid-plate space. Now it is time to consider the way in which this picture will change with frequency. Then, finally, these pictures will be useful for measuring instantaneous flow to the grid.

Figure 24 shows three different instantaneous appearances of the space charge when the signal is of *low frequency*. The grid is first at its most negative point, then at the average or bias value, and finally at the least negative point. There is a difference in the density of shading of the three diagrams to indicate the difference in the current and hence the density of moving electrons in the

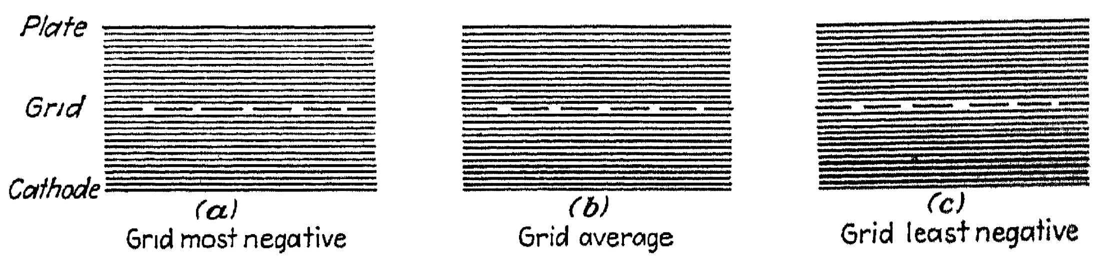

FIG. 24.—Instantaneous space-charge densities in a triode at low frequency.

interelectrode spaces. But there is no wave or bunching tendency to be noted. This is correct because the frequency is low. Any change in space charge caused by a given shift in grid voltage takes place in practically no time compared with the slow changes of the grid voltage.

From the instantaneous low-frequency pictures of Fig. 24 it is clear that there is induced in the grid at every instant the same magnitude of current from electrons traveling from cathode to grid as from electrons traveling from grid to plate. The two effects, though equal in magnitude, are, however, opposite in sign. Consider that the electrons approaching the grid are adding positive charge to the grid, while those leaving are removing it—all by induction, of course. Since the rate of passage of electrons from cathode to grid is equal to the rate of passage of electrons from grid to plate, none being collected by the grid, the zero total instantaneous induced current in the

grid is assured. (All these statements assume, of course, completely negligible transit time for the electrons, an assumption that is certainly valid for the lower frequencies.)

Next, in Fig. 25, is a case of fairly high frequency. This time it is again assumed that the electrons travel fast enough in comparison with the a.c. period so that no discernible grouping of the electrons into alternate regions of greater or less density occurs in the grid-plate region. The frequency chosen, however, is not so low as to permit disregard of the transit time in the cathode-grid region which is a lower velocity space. So now, when the three pictures of Fig. 25 are drawn for the three positions of instantaneous grid voltage, care must be taken to include

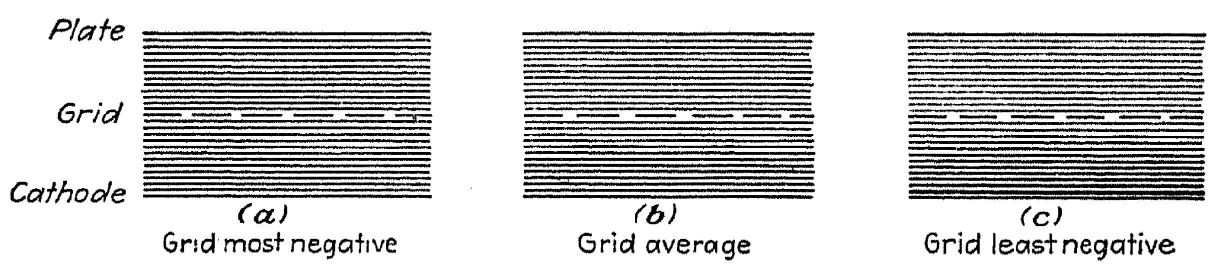

FIG. 25.—At high frequency, the electrons in the grid-anode space may be those that left the grid-cathode region a half cycle earlier.

this delay action. Thus, when the grid is in its least negative position (Fig. 25*c*), there will be the expected increase in density of electrons leaving the immediate cathode area; but as the reader looks away from the cathode toward the grid, he sees that at each point the instantaneous density will disclose more and more of the effects of the earlier instantaneous values of the grid voltage. Thus the density near the grid of Fig. 25*c* is low because the electrons there left when the grid was more negative than at the instant depicted.

Figure 25 has illustrated a special case. A transit time has been assumed that is precisely such that when the grid is least negative the electrons in the grid-plate space are those which left the cathode region a half cycle earlier, when the grid was most negative. This accounts for the lack of synchronism in changes of shading between the cathode-

grid and grid-plate regions from one instantaneous value of grid voltage to another. In particular, notice that now it can no longer be said that the effects of moving electrons on the two sides of the grid cancel one another. Now the rate of flow of electrons from cathode to grid is not equal to the rate of flow from grid to plate. The average values are admittedly the same because no electrons are being collected. But the instantaneous values may differ appreciably, and the grid now takes a.c. current.

It is also clear that the phase of the resultant grid current with respect to the a.c. grid voltage may have quite a range of magnitudes. It simply depends upon how large the transit time is in relation to the spacings, and it would truly be necessary now to integrate the separate effects of all the electrons to calculate precisely the magnitude and phase of the grid current. However, there is no reason why the grid current must be 90 deg. out of phase with the grid voltage, thus representing merely a change in the apparent capacitance of the grid due to the presence of electrons. The induced current in the grid may well have a component in phase with, or exactly 180 deg. out of phase with, the grid voltage.

If the application of a signal to the grid of a negatively biased triode results in its taking current, a component of which is in phase with the applied voltage, then energy is being absorbed at the grid. It is interesting to recognize where that energy goes. Closer study would show that the energy is used in imparting an average acceleration to the electron stream.

At the close of this chapter showing how electrons in motion produce currents, a problem will be considered that would arise in a typical tube intended only for operation at ultra-high frequency. Figure 26 shows the output end of one form of velocity-modulation tube whose input system was studied in the previous chapter. The electron stream may be supposed to have been velocity-modulated earlier along its length. Thus, when it flows from cylinder

A to cylinder *B*, an imaginary observer in the gap between cylinders will see a stream whose electron density is fluctuating in time. It is evident at once that when the electrons emerge from inside the cylindrical shield, the electric-flux lines that they produce will begin to transfer from one cylinder to the other. Thus an electron that goes out of *A* into *B* has transferred a positive charge by induction from *A* to *B* through the external circuit connecting them. It is clear, then, that output current can be taken from the beam by an external load or output impedance connected between two electrodes, neither of which collects any electrons. The beam may go right on

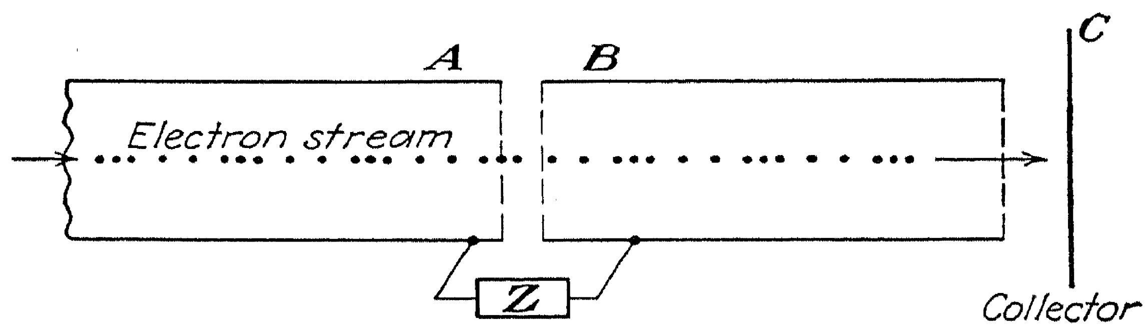

FIG. 26.—A current-density-modulated electron stream will produce a.c. currents in an impedance *Z* connected between the two cylinders through which the stream passes.

through *B* to collector *C*, where before collection it may actually be slowed down considerably to minimize plate-heating losses due to the bombardment energy of the electrons.

Centimeter-wave current flow, when due to a motion of electrons in free space, can take place between conductors neither of which is emitting or collecting electrons. The concept that might well be restated as a final summary of this chapter is that current flow in the external conducting system connecting electrodes of an electronic tube can always be correctly regarded as an induction effect due to the motion of the electrons in the space between electrodes. This is true for direct current, low, or ultra-high frequency.

CHAPTER VI

Time Delay in Electromagnetics Becomes Important for Microwaves

Elapsed Time.—It is common knowledge that "radio waves travel with the velocity of light." This statement acknowledges that if certain electrical effects are produced at one point, say a transmitting antenna, these effects apparently move out with a definite, finite velocity and are felt at some other point, the receiving antenna, at a later time. Time passes while the waves are in transit. If the generator were to change frequency or amplitude or drop out altogether, the receiver would not find out about it until a proper interval of time had elapsed.

This fact has a great deal to do with circuits, voltage, impedance, and the many factors that must be considered next in the study of the similarities and the differences between high and low frequencies. Unless the propagation or transmission of signals is being considered, what is the importance of radio waves? Certainly we have all worked correctly many circuit problems, with either lumped constant impedance or distributed constant impedance, without consciously including anywhere in our equations the statement that a lapse of time is required before effects are felt at a distance from the source. How fundamental is this delay effect? How necessary is an understanding of it? It is extremely important. Much of what goes on at microwave frequencies can be traced to this delay or, as it is ordinarily called, "retardation." Most of the difference between high and low frequency is a result of retardation; remove this effect from nature and there would be little reason to write this book. In fact, it would

be meaningless to ask what would happen to the laws of electricity and magnetism if retardation were excluded, because, except for direct current, retardation is required to make the other laws consistent. The reader will not yet have reason to agree with this statement unless he has had prior experience with the displacement-current concept discussed in Chap. VIII. It will suffice at this point to add that the retardation effect is simply one way to express the response of basic laws to variations in frequency from direct current on up.

Retardation.—The next point is just how to describe this phenomenon of retardation so that it is clear how it enters into all electrical effects. It is not sufficient to speak of a delay in the propagation of radio waves; at least not now, because the relation between waves and circuits and impedance have not yet been clarified. So we start at the very beginning. A charge of electricity has an electric field about it—a way of saying that it is ready to exert force on any other electric charge placed as a test in the field of the first. If the strength of the charge is altered in some way, the intensity of the surrounding field will follow in proportion. All this has been discussed already; nothing yet has brought in the effect of frequency. What must be added is a statement of how long it takes for the field intensity to change at each point in space when the source strength is changed.

Suppose the charge is suddenly increased. The increase will not show up in the surrounding space as an instant change in the field intensity there. As a matter of fact, at the first instant it will remain everywhere at its original value. Then, traveling outward from the charge at the velocity of light in that medium, the field increase will appear about the charge. As Fig. 27 shows, the front of the outgoing increase of electric field will be a sphere with the charge at its center, expanding in diameter so that a point on the surface of the sphere travels outward with the velocity of light. Every point in the surrounding medium

eventually attains the new value of electric field corresponding to the new value of the electric charge at the center, but there will be a delay in time between the change of charge and the change in the electric field it produces at each point. This delay will be just equal to the length of time it would take a messenger delivering information about the change to travel at the velocity of light from the charge to the point in question.

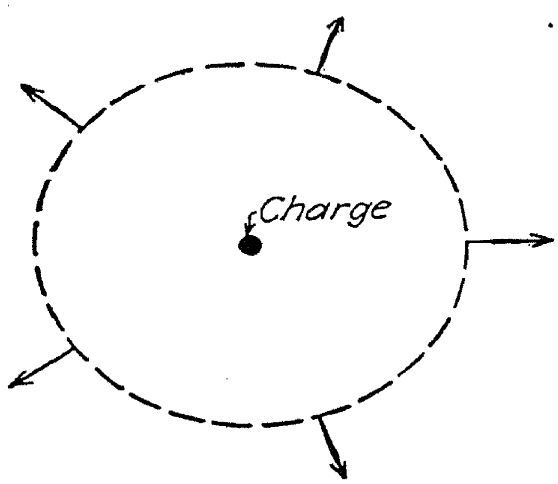

FIG. 27.

A similar thing happens to the magnetic field. If current is altered in a wire carrying current, the magnetic field that surrounds the wire will not change in phase with the current; it will lag behind, the time delay depending upon how far away the point where field is being evaluated is from the point where current flows.

Time Lag, a Universal Effect.—When charges and currents are changing in time, the effects they produce always lag behind those changes. Each point in the space that surrounds each charge and current element feels the change,

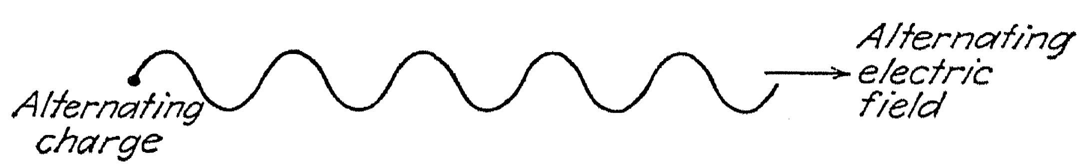

FIG. 28.—Each change in charge may be thought of as sending out a field in the form of waves.

one point after the other in turn in the order of their distance from the source. If a source is an alternating one (Fig. 28), then the effect of retardation and time delay can be demonstrated by a wave action. The idea of a wave's moving outward from the source is a correct and convenient one because it so nicely pictures each point in the path receiving the effect eventually, at a time dependent on its location and the wave velocity. Notice also how neatly the idea of a variation of instantaneous phase with space

is pictured by the idea of a traveling wave. Of course, this is a very simple case to illustrate the principle. In any practical case, the outgoing waves must come from a complexity of distributed sources, not a point source. Also, the outgoing waves will bounce against the boundaries of the region, *i.e.*, against conductors and dielectrics. The total electric field will come not only from the retarded action of electric charges but also from the inducing effect of changing magnetic fields caused by time-varying currents. Thus the resulting effect at any one point will, in general, be far different from the simple sine-wave action pictured. For the moment, it will be perfectly satisfactory to think of the simple time delay in all changing effects.

Inductance of a Circuit.—In all conventional low-frequency circuit work, the inductance of a circuit is thought of as being due to the rate of change of linked magnetic flux. The magnetic flux is assumed to be directly in phase with the current. The induced voltage, being then proportional to the rate of change of current, is 90 electrical degrees out of time phase with the current. This is perhaps the commonest notion in steady-state a.c. theory as it is learned by the electrical engineer—a reactance drop is 90 deg. out of phase, or in quadrature, with the current through the reactance. Yet considerations of the often neglected retardation effect have disclosed that the time delay between a current and any effect of that current depends on precisely how the sources are situated in space with respect to the location of the effect. It is beginning to be evident that if the frequency is so high that the current changes appreciably before its magnetic-field change begins to be noticed around the circuit, all our ordinary circuit ideas are in for a bit of transforming.

CHAPTER VII

Retardation and Radiation Influence Basic Circuit Behavior at Ultra-high Frequency

The reader has been prepared to delve into centimeter-wave circuits and transmission lines and antennas by the introduction of the retardation effect. It is basically a simple matter that there is a space-travel time delay between currents or charges and their effects. Also, it is easy to appreciate why retardation is of negligible importance if the system is physically small compared with the wavelength. If retardation is now included for a correct consideration of the operation of ultra-high-frequency systems, the direct effect and the implications are not quite so easily written down. The remainder of the text is required to assemble the most important of the concepts that stem largely from this retardation effect and that control characteristics of microwave systems.

Starting with the simplest circuit considerations, the reader will discover what the addition of retardation will do to his notions of what a circuit is and how it behaves. At the very beginning, it is fitting to question whether there is such a thing as a circuit at these extremely high frequencies. Perhaps there are only fields and waves, and words like "lumped circuit" and "circuit constant" have little meaning. If this be true, the change can hardly take place suddenly as the frequency increases. It seems more reasonable to expect a gradual extension of or departure from circuit concepts in passing from the lower frequencies to the very high ones. In any case, the effects and their various influences can be seen if a simple, single loop of wire (Fig. 29) is excited by an alternating voltage and

the action is followed as the frequency rises. For certain reasons of convenience that will be evident later, the shape of the loop has been chosen approximately rectangular with one dimension quite large compared with the other.

Simple Circuit at Low Frequency.—At low frequency, the analysis of this problem is familiar; the applied voltage is used up in IR and IX drops. Again, for convenience, without interfering with a demonstration of principles,

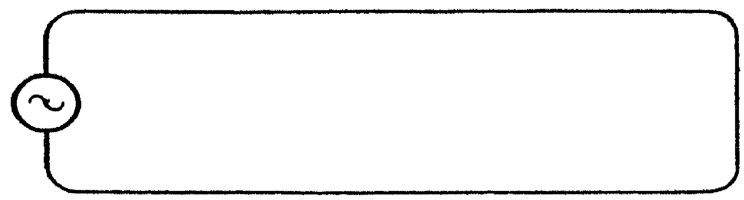

Fig. 29.

a perfect conductor is assumed. There will then be no ohmic drop. It is recalled that the remaining IX or reactance drop comes from the induced voltage, which in turn is a result of the rate of change of magnetic flux produced by the current flow. The current flow everywhere around the loop is the same in both magnitude and phase. The magnetic flux, retardation being neglected at low frequency, is in phase with the current that causes it. It is

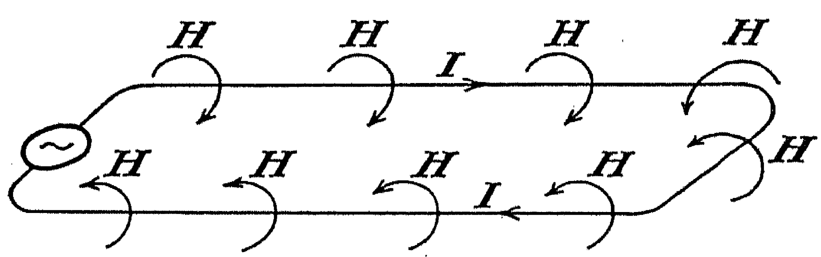

Fig. 30.—The low-frequency picture includes a uniform loop current surrounded by a magnetic field everywhere in phase with the current.

linking the current somewhat as pictured in Fig. 30. The region of the loop, then, is surrounded by a magnetic field that is everywhere of like phase.

To complete the picture of what takes place around this simple circuit at low frequency, electric-field distribution must next be considered. When a voltage is applied across the terminals of the loop, the integral of the electric field directly across those terminals is set by the magnitude

and phase of that applied voltage. To determine just how much electric field, in both magnitude and phase, will exist between parts of the loop other than those directly across the generator terminals requires a little more study. It seems reasonable to suppose that the final electric-field distribution will be somewhat as pictured in Fig. 31. Here the electric field, first across the terminals and then between the two long sides of the loop, gradually grows weaker as the end of the loop is approached. Finally, between the two ends of the wire (which actually touch to form a complete loop) there will be no electric field, as the very evident short circuit existing there requires.

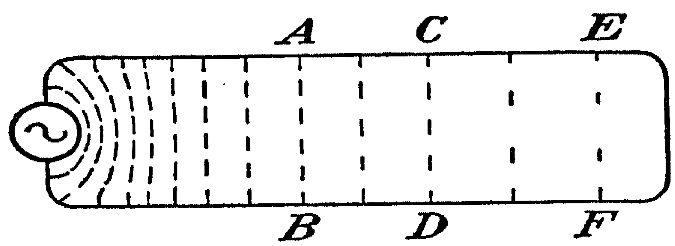

FIG. 31.—The dotted lines represent electric field which gradually loses strength from the generator to the end of the loop.

It is now possible to check whether this distribution is correct by considering what it is that determines how the electric field shall decrease from its magnitude at the input terminals. For example, how does the electric field across AB differ from the electric field across CD in Fig. 31? It is evident immediately from Faraday's law that, since in the rectangular loop $ABDCA$ there is a certain rate of change of magnetic flux, an induced e.m.f. will exist. In other words, the reason that the voltage difference across the loop at AB is not the same as that across CD is that some voltage is induced by the magnetic flux between the wires. It is already known that the induced voltage is of such phase as to buck the applied voltage. This means that the voltage difference CD is smaller than the voltage across AB by just the amount contributed by the changing magnetic flux linked by the rectangle $ABDCA$. In a similar way, voltage EF will be

somewhat smaller than voltage CD. The vector diagram of Fig. 32 appears to describe the situation completely. This diagram shows the current I, in phase all around the loop, as the reference vector, and the applied voltage 90 deg. ahead of the current in phase, as is fitting for a pure reactance circuit. In addition, a series of short vectors is drawn on this diagram representing the voltage differences around the loops $ABDCA$, $CDFEC$, etc. Each of these small differential voltages subtracts from the applied voltage until finally at the end of the loop the induced voltage has completely subtracted from or neutralized the applied voltage, and there is no voltage difference remaining across the conductors of the loop. One way to look at this, of course, is that the current in the loop will adjust itself in magnitude so that the amount of the net voltage resulting from the changing magnetic flux caused by the current will exactly cancel the applied voltage, as indicated in the diagram.

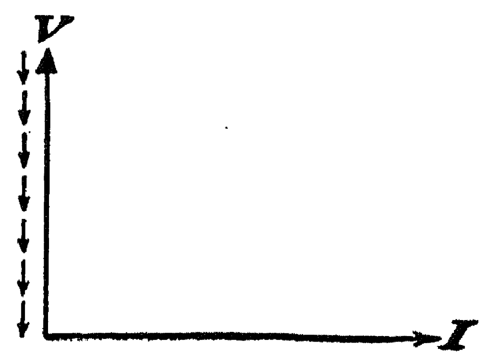

FIG. 32.—A voltage-current diagram for the low-frequency circuit of Fig. 30.

Distribution of Charges.—It can next be noted that in this low-frequency picture, even though retardation was ignored, another approximation not yet mentioned seems to have been introduced quite by accident. If there is an electric field between the two long conductors that together essentially form the loop of the preceding diagrams, then, of course, the electric lines that end on the conductors must find electric charges there. Notice that the analysis so far has at least been consistent in that with a perfect conductor there can be no tangential electric field along the conductor. If there were, it would cause an infinite current flow. Thus it is a satisfactory result that the electric lines of force will come in perpendicular, much as shown in Fig. 31. But there is then a distribution of positive and negative charges on the loop, as indicated in Fig. 33. Near the generator end the charge density along the conductors

will be more dense than near the end of the loop because there is a greater amount of electric field there.

If there is a distributed charge density along the loop, then it follows that as the electricity goes through its alternations, the charges on the conductors will oscillate from positive to negative. The picture of Fig. 33 is, in other words, simply an instantaneous picture. For these

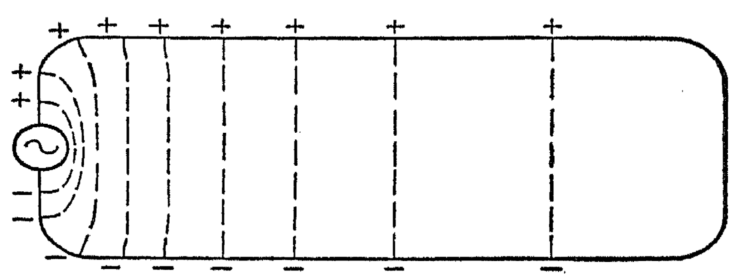

FIG. 33.—Charge must always be distributed over the loop, even at low frequency.

charges that rest on the conductors to oscillate in magnitude from peak positive to peak negative during the cycle, it will be essential that there be some charging current flowing from the generator to the loop. Perhaps it will be more convenient at this point simply to note Fig. 34, which recognizes these capacitive effects as the result of distributed capacitance existing across the loop, an effect often neglected at low frequency.

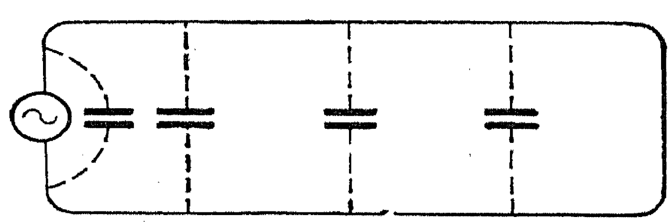

FIG. 34.—Charging current effects are represented by the distributed capacitance.

This neglect of the stray capacitance may be completely justified. The point is brought up here, not for fear that insufficient accuracy will be obtained otherwise, but simply as a matter of principle. When the complete picture of electric and magnetic field linking the loop is considered, it must be noted first that there is always a component of current flow to those distributed capacitors. Secondly, it must be noted that, since some current will go into these

capacitors from the generator, the total current flow around the loop cannot be exactly uniform. There will always be some current at each point that flows through the parallel capacitive path and fails to go completely around the loop. This is pictured in Fig. 35, which demonstrates that even at low frequency, and without mentioning retardation as a factor in determining current distribution around the loop, the current flow cannot be exactly the same all the way around the loop. It must change at each point by the differential amount necessary to charge the distributed capacitors.

For the low-frequency picture, at any rate, it can always be assumed that the wire size is so small compared with the

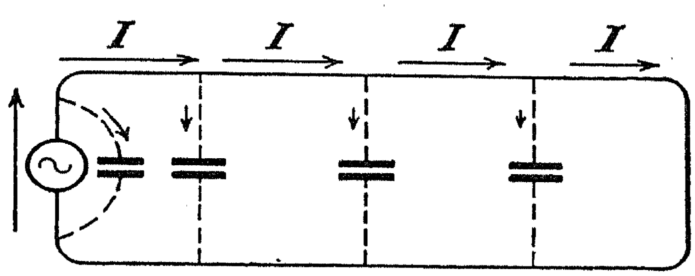

FIG. 35.—Even at low frequency the loop current must be nonuniform, though perhaps very slightly.

size of the loop that the amount of charge that must exist on the wire for the required electric field between various parts of the loop will be trivially small. Thus the approximate low-frequency picture of the performance of this basic circuit may be summarized as follows: As a result of the applied voltage, there is a certain current flow in the loop. That current, for a perfect conductor, lags the applied voltage by 90 deg., and is the same in magnitude and phase all around the loop. The loop is surrounded by an electric and a magnetic field. The magnetic field links the loop's current and is everywhere in phase with the current, *i.e.*, 90 deg. out of phase with the applied voltage. The rate of change of magnetic field, due to the loop current, causes an induced e.m.f. that succeeds bit by bit along the loop in bucking out the applied voltage. There is, in addition, an electric field in the region of the loop that

exists across different parts of it. This electric field is fixed as to its integral directly across the generator terminals by the amount of the voltage applied, and it falls off to zero as it progresses down to the end of the loop, all of this being determined by the linking magnetic flux that is changing in time.

Effects of Retardation on Basic One-loop Circuit.—It is time now to add retardation and to consider what that will do to the pictures drawn up on the behavior of the basic one-loop circuit. First of all, let it be agreed that the results already obtained will stand unless a reason is seen to eliminate them. Now, let there be in the loop a current flow that results from the applied voltage. That current flow may be tentatively said to be the same all the way around the loop. This is making a little more drastic assumption about the distributed capacitance effects than at low frequency. However, it will be helpful to keep the distributed capacitance effect out of the picture in order that the retardation effect not yet considered may have the whole stage.

Suppose, in other words, that though the cross section of the wire remains very small, the length of the loop becomes appreciable compared with wavelength. Let it not be too long, at first, but at least long enough so that an appreciable time delay may be expected from one end of the loop to the other. Consider, first, the magnetic flux near the generator end. The magnitude of magnetic flux and the phase of the flux near the generator contributed by the current in its immediate vicinity will be very little different from the low-frequency case, but the contribution to flux near the generator from the current some distance away from the generator terminals will arrive late. The resultant total magnetic flux at the generator will be somewhat behind the loop's current in phase, it seems. Consider next a point near the center of the loop. Here, again, a good part of the magnetic flux is contributed from the two ends of the loop, or at least

by the current between the ends and the center. Again, there will be some phase delay between the total magnetic flux at this point and the current flow in the loop. The situation at the far end of the loop is much the same as that at the generator terminals.

Now draw another vector diagram showing how the electric field falls off between conductors of the loop. Once more, until there is some reason to change, it is assumed that the electric-field distribution is the same as at low frequency. So Fig. 36 shows an applied voltage and a loop current flow that is 90 deg. lagging with respect to that voltage. Now the various differential e.m.fs. due to

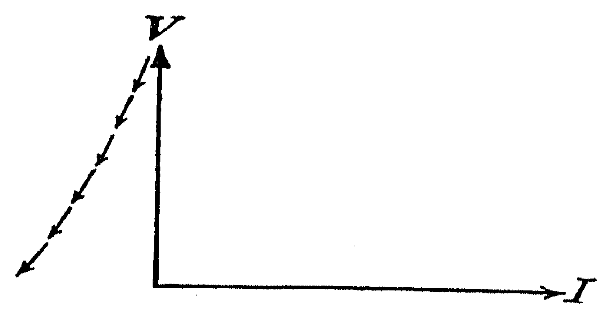

FIG. 36.—Assuming retardation and a uniform loop current I with a 90-deg. leading voltage V, the vector diagram refuses to balance.

the flux linking the small rectangles like $ABDCA$ (Fig. 31) must be subtracted much as before. These differential induced voltages, being 90 deg. out of phase with the magnetic flux, are not exactly 180 deg. out of phase with the applied voltage. An additional phase angle between the magnetic flux and the current has come in because of retardation and must be added to this total angle. Thus this differential voltage vector is pictured with a slight angle with respect to the applied voltage as shown in Fig. 36. For convenience, each differential voltage vector in the diagram is shown with about the same angle introduced by retardation; these may actually differ, but at this point the reader is obviously interested in qualitative rather than quantitative effects.

Notice then how a rather fundamental change in the apparent characteristics of the circuit presents itself, for as the differential voltages that result from the changing

magnetic field of the current are added down the loop, it is found that each of them is always somewhat more than 90 deg. out of phase with the current. When the end of the loop is reached, the resultant induced voltage fails to subtract out, or compensate for, the applied voltage. The system simply will not balance, and some of the assumptions must be revised.

One thing that can be done and that would fit all the facts so far is to remove the initial assumption that the applied voltage and the current are 90 deg. out of time phase. For example, the current vector shown in Fig. 37 might be drawn again as a first step. This time, however, add up each of these induced voltages due to the current until finally, when the resultant vector OP is reached, it is recognized that this amount of voltage in exactly opposite phase must be applied by the generator in order to compensate exactly for the reactance drop of the loop. The only thing about this that takes special attention now is that in the end there are in the loop a voltage and a current that are not 90 deg. out of phase. This means, of course, that there is a dissipative component in the volt-amperes being supplied by the generator. The question immediately comes up, then, as to where the energy goes. There is an energy loss, this reasoning seems to indicate, that is a direct result of taking retardation into account. The energy presumably cannot go into ohmic losses in the loop because a perfect conductor with no resistance whatsoever has been assumed. The energy must, if it is truly being lost, be *radiated; i.e.*, it is escaping in leakage of some kind in the electromagnetic field.

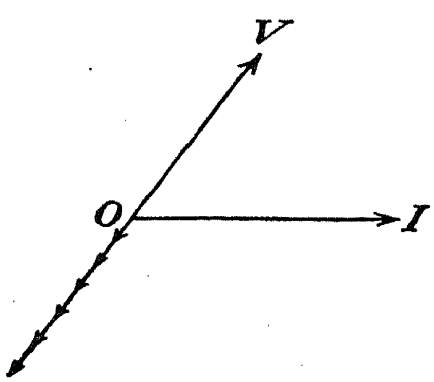

FIG. 37.—The diagram balances if I and V are not out of phase by 90 deg.

Of course, the initial assumption can be revised in one other way. It can be said that the capacitive effect should not have been neglected; the current flow is actually not uniform in phase and magnitude all around the loop.

In this way the initial analysis of the phase of the magnetic field at each point with respect to the current would be revised. With such a postulate, the simple current reference vectors of Figs. 36 and 37 could no longer be used because there would not be any one single current everywhere in the loop. But it is conceivable that, even though the current is now assumed to change in phase and magnitude about the loop, the current flowing into the loop right at the generator terminals might come out exactly 90 deg. out of phase with the voltage at those terminals, just as in the low-frequency case.

Now it was not the intention of this discussion to make possible at this point a complete understanding of the situation that has arisen. This is what was meant when it was stated that the rest of the text would be required for the analysis of microwave concepts. But one thing has been shown—that the added effect of retardation, usually justifiably neglected because the frequency is low, or rather because the circuit is small in dimensions compared with wavelength, is capable of making some marked changes in concepts of the circuit. There remains the choice of at least one if not both of the following conclusions:

1. In a simple one-loop circuit that has dimensions comparable with wavelength, the current flow is not uniform in magnitude and phase all the way around the loop.

2. Retardation brings in a kind of energy loss in the loop by some kind of electromagnetic-field leakage that results in the applied voltage's being other than 90 deg. out of phase with the current.

If either of the two situations exists—and it appears that one or the other, or both, *must* exist—then several marked changes are immediately noticeable at high frequency when it is viewed from the usual low-frequency-circuit concepts. The inductance effect that comes from the time-varying magnetic field is not necessarily reactive, or perhaps it should be said that reactance is not necessarily completely loss-free. Radiation is an important concept in microwave circuit behavior, a direct result of retardation.

CHAPTER VIII

Displacement Current Is as Effective as More Common Current Flow

The preceding chapter furnished us with an impression of the complexity that can be unearthed when even a simple circuit is investigated at microwave frequencies. We should not think this is a necessary situation at these frequencies because of the way this one example went. It happened to be an exceedingly simple circuit at low frequency, at least, as to current distribution. For various simple forms of the loop it would have continued to be simple even as to details at low frequency; we could have gone on to find the inductance of the loop and the ratio of current to voltage. In contrast, of all the problems that lend themselves to analysis at ultra-high frequency, those involving loops of wire are among the most difficult. It happens that one of the simplest problems at low frequency (and one that might be expected to be a common example taken to illustrate broad principles there) is at the same time one of the most difficult to understand clearly at ultra-high frequencies. Also, and not for the same reason, certainly, it happens to be true that the loop-of-wire type of circuit is the commonest circuit at low frequency and far from the commonest at centimeter wavelengths.

But this anticipates the discussion to come. For the moment, it should be pointed out that it is difficult to generalize about ultra-high frequency by considering one simple example. We are not too sure, for instance, what is the best way to approach even that one-loop circuit when the wavelength is comparable to the loop dimensions, but at least it is possible to see what has to be studied

further. Obviously we must understand better fields and waves and their relation to currents on the conducting sources and boundaries of the region, to voltage of the generator, and to power flow by radiation. Further consideration of general principles simply cannot be avoided.

The present situation with regard to acquiring a feel for microwave concepts may appear to the reader to be as follows: To describe microwave setups with any success at all means that physical laws of a rather general nature must be introduced, and they must be gone over a bit so that they will appear wholly reasonable. To dwell on these general theories, however, is not so pleasant for most readers, perhaps, as it would be to maintain a closer tie at every moment with a centimeter-wave phenomenon. Thus it is well to be convinced that the general laws are necessary and worth studying. This conviction is obtained when the analysis of a simple loop is attempted, and so many interesting and new conclusions appear. At least, so the author hopes; for this chapter is indeed a study of general theory, applicable at any frequency, but often not familiar to those who have worked mainly with the lower frequencies. The theory of this chapter should emerge in the form of certain clear concepts that describe the behavior of electromagnetic fields. With them, the understanding of microwaves is assured; every hour of further study will build on a firm base. Without these concepts, time spent in trying to understand ultra-high frequency may be dissipated in forever seeking a base from which to launch the attack.

First of all, we want to look for concepts that lend themselves to easy, accurate physical pictures that can be employed appropriately anywhere in the frequency range. *Displacement current* is one of these concepts. It is not a necessary concept. The equations of electrical theory could be written down correctly and solved without ever referring to one of the terms as a new kind of current. But displacement current does a good deal to build up just

such physical pictures as we are seeking. The displacement-current concept attaches the notion of a current to the rate of change of electric displacement flux.

Displacement Flux.—So far, this text has not mentioned displacement flux, although it has mentioned electric-field lines whose density in space is a measure of the force a unit test charge would experience. The electric displacement flux is simply the dielectric constant times the electric-field strength. It is analogous to magnetic flux, which is the product of permeability and magnetic-field intensity. The concept of displacement current shows that, if the electric field is changing with time in some region of space, the effect of such a change is just as though an alternating current (of the type that flows in a wire, or as free electrons between electrodes of a vacuum tube) were flowing in the direction of the field. Exactly what is meant by the statement that, if the electric displacement is changing at some point, it is as though there were some current density there? The meaning is that this displacement current produces an alternating magnetic field and nothing else. It is not claimed that a density of charge is suddenly created in space by an electric field and that it moves to form a current; only that if there is a changing electric field it is accompanied by a change in magnetic field. The magnetic field is of such strength and direction and distribution as would be predicted by considering the changing electric fields as a distributed alternating current whose density everywhere is proportional to the rate of change of the electric field or displacement flux.

This displacement-current idea will be used to explain resonant cavities, the circuits of microwaves, which will be the subject of the next chapter. But first we must justify the concept and see how reasonable and lacking in mystery it really is. We start with an elementary picture, one that is not special but general, because complex cases can be built up with the background gained from the elementary example to be considered here. Figure 38

shows a linear current that is alternating and setting up a surrounding magnetic field, a part of which is sketched in the figure for one instant of time. The magnetic field pattern pictured is taken in a plane perpendicular to the current flow, and the linear current element may be part of a more complex system of currents in conductors. All the conclusions drawn will be true for the whole as well as the part. That is why no generality in principles is lost by consideration of the elementary case.

Induced Electromotive Force.—Notice the rectangle sketched with dotted lines in Fig. 38. Through it passes

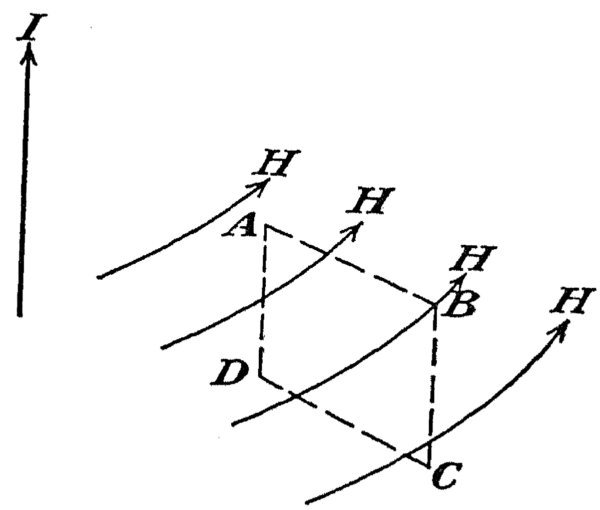

FIG. 38.—A rectangular loop *ABCDA* whose plane passes through *I* is linked by magnetic field *H*.

changing magnetic flux caused by the current element, so we conclude that there is an induced e.m.f. around this rectangular loop. Admittedly, when we do this, much as in the previous chapter, we are quoting the original experiments of Faraday on the induction of voltage in a circuit caused by changing magnetic flux linking the circuit. The reader is certainly justified in questioning whether these experiments and the conclusions from them really apply at ultra-high frequency, particularly in view of the discussion in the text up to this point. The answer is that if we assume that Faraday's law does indeed extend to cover the whole frequency range, and into regions where there is no conducting circuit, we shall be led to certain conclusions that have been verified experimentally.

It is taken, then, as a general law applicable to the whole frequency range, that around any loop that might be drawn in space there will be an induced e.m.f. proportional to the rate of change of magnetic flux linking that loop.

Accordingly, if there is *no* electric field along the line *AD* of the rectangular loop—and only magnetic field has been said to be present—then there must be some electric field along the remainder of the loop. This conclusion is drawn because there is a rate of change in magnetic flux through the rectangular loop *ABCDA*. The difference between the field along *AD* and along *BC* must be accounted for by that induced e.m.f. (since, from symmetry, the e.m.f. along *AB* and *CD* will cancel). Thus the entire magnetic field about the linear current can be explored, and it can be concluded that there must be electric field throughout the surrounding space. This electric field is alternating and is associated in strength with the rate of change of magnetic field.

Now obviously this conclusion would be reached with or without the inclusion of retardation. Even neglecting retardation, the magnetic field would be linking the current; and the rate of change of the magnetic field would imply the presence of a changing electric field in the surrounding space. One general conclusion of completely wide applicability may then be noted at this point: Whenever there is a time-changing magnetic field in space, there must also be an accompanying time-changing electric field in that space.

But retardation is responsible for the magnetic field's reversing in direction with distance as points away from the current element are considered at any instant; *i.e.*, points distant from the conductor will lag behind in oscillation phase. Retardation is responsible also for the electric field's doing the same thing. This is seen as follows: Through one rectangular loop that might be drawn in one part of the region (Fig. 39), the time rate of change of magnetic flux (indicated diagrammatically by

the wave) is such as to cause an increase of electric field with distance from the element. In another region, say farther away, the reversal of the magnetic flux causes a decrease with outward distance. Since the electric field has no d.c. component and yet oscillates in the sign of its increment with distance, it follows that the electric field

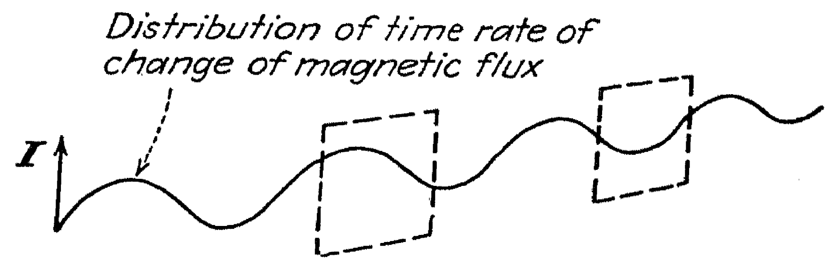

FIG. 39.

will show periodic reversals with distance in its space pattern at every instant. These reversals will be identical, in their number per unit distance, with those exhibited by the magnetic field. The wavelength of each, in other words, will be the same. This instantaneous picture is indicated in Fig. 40. If more time were spent in studying this diagram (and if successive diagrams were drawn to

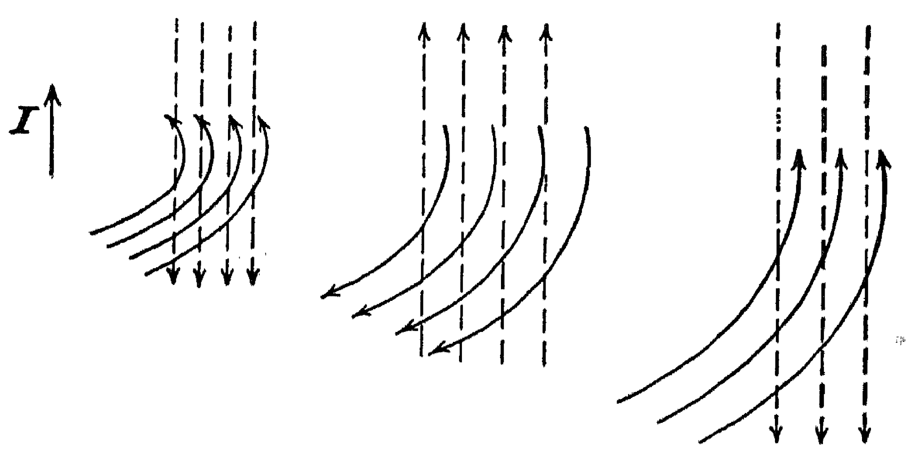

FIG. 40.—The magnetic field lines (solid) oscillate in direction with distance from I and are accompanied by electric field lines (dotted) which do likewise.

disclose the change of the pattern with time), it would be fairly easy to see that the waves of electric and magnetic fields are in phase everywhere. They have their maxima and minima together in time at each point of space.

While this picture of the spreading wave of magnetic field with its accompanying wave of electric field is still

in mind, a second situation can be considered. Figure 41 shows two conducting spheres that are being charged by an alternating voltage generator. The dotted lines are intended to represent crudely the distribution of electric-field lines over part of the region that surrounds the charges. These electric-field lines pictured are contributed by the charges with retardation considered. Obviously no attempt at precision has been made in this diagram, yet

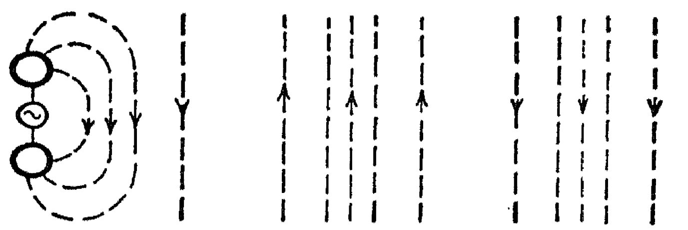

FIG. 41.—The electric field caused by the charges on the spheres oscillates in direction with distance at every instant.

the idea is brought out that retardation enters to cause periodic reversals of the electric field as distance from the charges is changed. And so now it can be concluded that there must be present in this field a time-varying magnetic field, a situation completely analogous to that pictured in Fig. 40. The proof is easy: If the electric field oscillates with distance at every instant, then a little Faraday loop

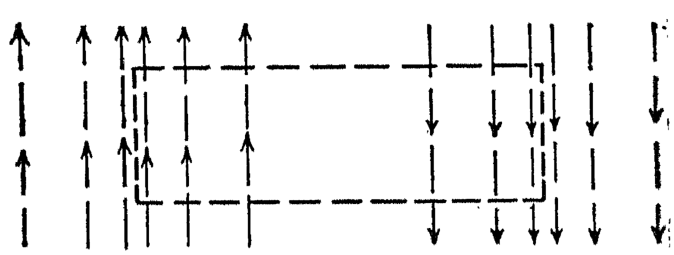

FIG. 42.

(Fig. 42) may again be drawn, and the result of integrating around it, wherever it may be placed, is an e.m.f. that is not zero. According to the law stated above, if there is a net e.m.f. around a loop in space, there must be a changing magnetic field linking that loop.

Again, the wavelengths of the magnetic field and electric field will be the same, and both quantities will have their maxima and zeros at the same place and time.

At this point, the concept of displacement current may well be introduced to make the analysis tie together easily. Whenever at any instant a magnetic field is not spread uniformly through space, it is possible to explore the magnetic field by drawing small loops around which

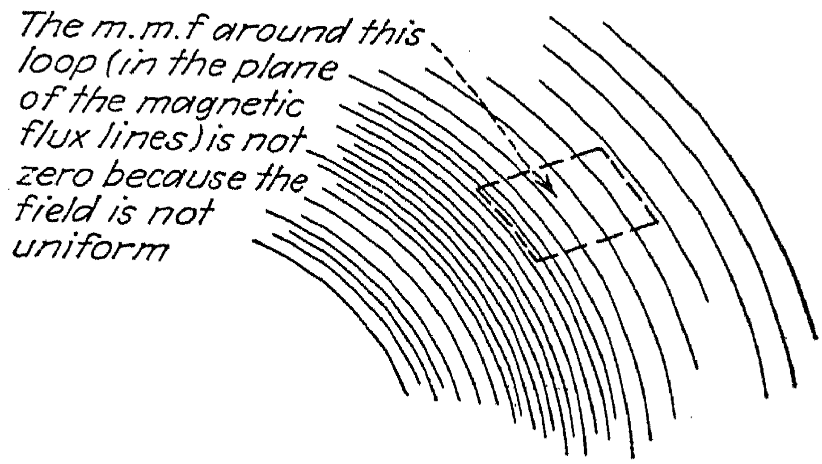

FIG. 43.

will be integrated not the electric field, to find the induced e.m.f., but rather the magnetic field, to find the induced m.m.f. This is done in Fig. 43. A rectangular loop has been drawn around which will be taken the integral of the magnetic-field intensity. That integral will, of course, be something other than zero because the magnetic field

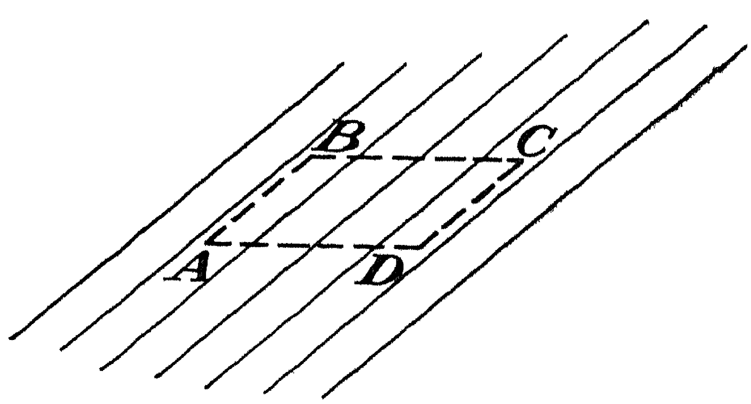

FIG. 44.—A portion of a uniform magnetic field.

varies throughout the region. If the magnetic field were uniform (Fig. 44), the integral would be zero; the only contributions, along the lines AB and CD of the loop, would cancel. Since the magnetic-field intensity may be derived by a consideration of the distribution of time-

changing electric flux, the electric flux's changing with time may be regarded as responsible for the magnetic field's presence. This shows the reasonableness of the concept that the magnetic field is caused by a displacement current. The existence of changing electric flux in space produced encircling m.m.f. just as surely as though there existed an ordinary everyday conduction current in the space with a current density everywhere proportional to the rate of change of electric field. Thus it could have been concluded that in a space excited by oscillating electric charge there must be a magnetic field by the simple process of endowing the time-changing electric displacement flux in the space with the magnetizing ability of a current flow.

Now here are two laws side by side in perfect symmetry, with convenient physical pictures, and with great potency: Changing magnetic flux in space gives rise to changing electric fields that tend to link the magnetic-flux lines. Also, changing electric flux in space gives rise to changing magnetic fields that tend to link the electric-flux lines.

This discussion may appear to have been devoted so far more to the idea of bringing out the relation between magnetic and electrical phenomena in space than to displacement current. But the reader may be assured that the discussion that brings in the idea of magnetic effects' being induced by time-varying electrical flux is of prime importance in understanding high frequency. We are all familiar, it is supposed, with the reciprocal idea that time-varying magnetic flux induces electrical effects. This is essentially Faraday's law, and most readers will only have to make the simple extension, if any, to the use of this law liberally in dielectrics as well as in conductors. The magnetic effects of displacement current may be expected, however, to be new to many readers, despite the fact that displacement current is not limited to ultra-high frequency.

It must be recalled that displacement current is proportional to the rate of change of the electric displacement flux. For direct current, the rate is zero. There is no mag-

netic field set up by a static electric field. With a given electric-field strength, the magnetic effects that do arise increase directly with frequency. As a matter of practical usefulness, the frequency generally has to be quite high before the displacement current reaches such proportions that the magnetic flux associated with it is noticeable. Thus the fact that a time-varying electric field acts like a current in giving rise to magnetic fields might be expected to be overlooked at low frequency where it will most often be masked by other effects. However, it is to be emphasized that the general law holds over the entire frequency range; changing magnetic fields give rise to electric fields, and changing electric fields give rise to magnetic fields.

Displacement current has been introduced by way of a general discussion of fields and waves in space. Once it is granted that displacement current can be used with complete confidence that the right answer will be obtained, we can complete more satisfactorily some pictures of a.c. circuits that would otherwise have bothered us considerably. In fact, as we pick up the concept of displacement current and apply it to many familiar cases, we are sooner or later struck with the thought that we should have appreciated it before. Consider the simple loop of Fig. 45*a*, which shows a generator tied to a condenser. Assume a low frequency and suppose that the distance across the generator terminals is small, with the condenser size also small. However, the condenser does break the circuit, *i.e.*, the conduction current. No actual passage of charge takes place between the plates of the condenser.

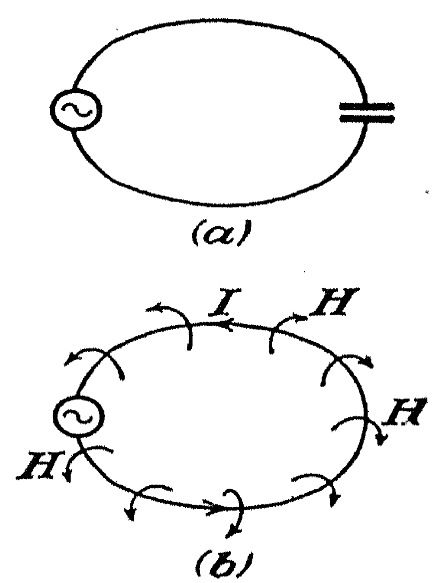

FIG. 45.

As for the magnetic field linking the loop, with a perfectly continuous loop such as that of Fig. 45*b*, there would be no difficulty whatever. The magnetic flux is con-

ceived as linking the current and having a strength dictated by the strength of the current. Suppose that the condenser's gap is extended as pictured in Fig. 46 so that now the condenser is an appreciable part of the loop. The dielectric constant of the material in the condenser can conveniently be made very high so that the displacement flux inside the condenser will be restrained along the path that would otherwise have been occupied by the current of the loop. What now is the distribution of magnetic field encircling the loop? Consider the magnetic-field lines like *abcd* as shown in the diagram by dotted lines. Do such lines even exist? Certainly there is no current passing along the dielectric if by current is meant only that which is associated with the motion of actual charges. Experimentally, however, the magnetic-flux distribution would be found to be the same in Fig. 45 as in Fig. 46, and the reason lies in the fundamentals that have just been reviewed. Probably the easiest way to imagine the form of the magnetic field linking the loop around the region of the condenser, where the current seems to be nonexistent, is to use the notion of displacement current. If there is a change of electric flux in that region, then that is the equivalent of common electric current as far as its ability to produce magnetic flux is concerned. The actual numbers that will be used in figuring displacement current between condenser plates will be such that the displacement current will precisely equal the amount of current flowing in the conductors of the loop. Thus the current is continuous, and the displacement current flowing in the condenser is actually equal in magnitude to the charging current of the condenser.

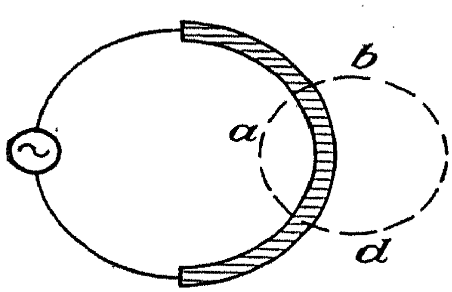

FIG. 46.

Now it can be said that the magnetic-field distribution of Fig. 45*a*, which shows the original small condenser in the circuit, is essentially the same as that of Fig. 45*b* and

Fig. 46. (There is always the slight distortion due to the area of the plates not being the same as the wire cross section and the condenser's end effects, but the present discussion is quite apart from such items.)

The identity of these magnetic-field distributions is precisely what the electrical engineer correctly uses all the time. For he would look at the circuit as containing inductance and capacitance in series and calculate the inductance, a function of the magnetic-field distribution, without worry over the major break in conduction current in the circuit. He always regards the current, in other words, as continuous. Therefore, he is attributing to the condenser's gap the ability to produce magnetic effects

FIG. 47.—A summary of Maxwell's equations for free space: changing electric flux produces linking m.m.f.; changing magnetic flux produces linking e.m.f.

with force equal to that of the conduction current that flows in the charging leads.

Further development of this picture and an appreciation of the significance of displacement current as a concept applicable to the entire frequency range will be obtained as we go on through the study of examples in the remainder of the text. For the present, it will be well to look at the two diagrams of Fig. 47, which contain a summary of this chapter. These show changing electric field with magnetic-field linkages and also changing magnetic field with electric field linking it. Appreciation of the facts that these two basic conditions can take place in space and that all field distributions at high frequency are a result of a superposition of these elementary effects will lead to an understanding of ultra-high-frequency circuits, transmission lines, and radiators.

Maxwell's Equations.—One more word is fitting, to aid many readers when they refer to other sources and also to give credit to Maxwell. His field equations applied to free space are a mathematical statement of the pictures of Fig. 47. When it is said that Maxwell's equations alone need be studied for a complete understanding of microwave phenomena, it is not much of an exaggeration. These equations predict the wave characteristics of electricity, state what the wave velocity is, and give the relations among the fields, the charges, and the currents of the system. It is no wonder, then, that a discussion of the mathematical theory of microwave systems usually starts from Maxwell's equations or an equation derivable from them.

An interesting lesson comes also from recognition of the fact that Maxwell wrote his equations before radio waves were known. This text starts from the known existence of electromagnetic waves and arrives finally at Maxwell's equations. That waves can be made to exist is known to every reader, and thus this text has begun with known and acceptable things. But the more usual approach, and probably the best one for a more thorough study than this, is to follow Maxwell. First, his equations would be made to seem reasonable and necessary as a statement of experimentally determined laws. Then, from those laws, the possible existence of waves would be predicted. The order of explanation is perhaps only academic, but it should not be forgotten that Maxwell predicted electromagnetic-wave phenomena.

CHAPTER IX

A Resonant Cavity Is a Self-enclosed Circuit

Equipped with some all-frequency laws concerning electromagnetic fields and, in particular, with some knowledge about displacement current, we are now ready to continue the study of ultra-high-frequency circuits. This time we can take a long step; we shall endeavor to investigate and understand reasonably thoroughly the commonest circuit, or the equivalent of a circuit at centimeter wavelengths. Because retardation must not be neglected at these wavelengths, radiation may be so great as to destroy the value of many otherwise excellent circuits. Circuits are nevertheless needed that will present high impedance between the terminals of an electronic tube, to pick up and transfer energy in certain frequency bands or absorb and block other frequencies. Every property of a circuit in low-frequency radio that makes circuits desirable is a useful property to possess at ultra-high frequency.

The way in which these requirements of usefulness are met in the centimeter-wave bands essentially eliminates radiation as a controlling factor. Radiation is outflanked by proper application of fundamentals already discussed. Thus it is not necessary to study this leakage of energy through electromagnetic waves any further at this point. In the important fact that centimeter waves will not appreciably penetrate a conductor is found the answer to the question of how to design circuits at these high frequencies so that the leakage of energy by radiation will not be excessive. If a circuit is to be built, it should be so constructed that the current flow around that circuit is self-enclosing. The practical realization of this trend

of thought is the common centimeter-wave circuit known as a "resonant cavity."

In Fig. 48 an attempt is made to picture a progression from the ordinary circuit to a simple self-enclosing resonant type of circuit. The diagram shows perhaps the most common impedance group in radio engineering, a tuned parallel circuit, the inductance tuning the condenser to resonance and yielding a very high input impedance across the parallel combination. Such high impedances between two points are needed in the centimeter-wave frequency band, but it is necessary that radiation be held to a small amount. Thus let us start out with two plates for the condenser, first closing those plates by a one-turn induct-

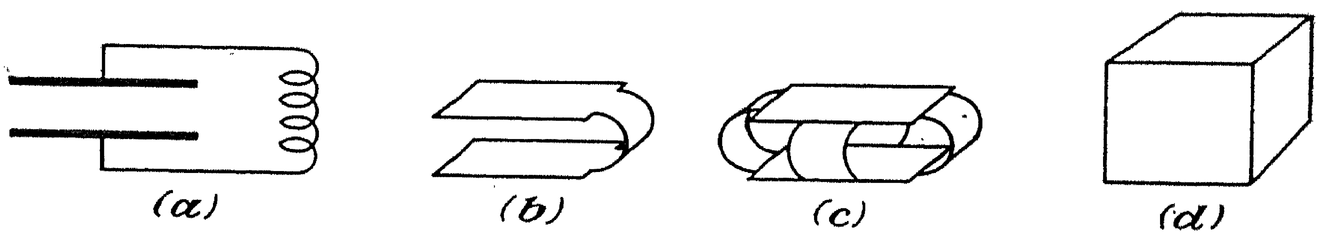

FIG. 48.—A progression from a common parallel LC circuit to a resonant cavity.

ance; this small inductance ensures that the circuit will resonate at an extremely high frequency.

Next, several of these one-turn inductances are placed around the condenser-plate edges. If desired, we may think of this arrangement as many small inductance bands in parallel. At any rate, the capacitor consisting of the two plates is closed by the surface on which current flows and is closed all the way around. Here is simply a closed box, and inside is the centimeter wave. This box may be stimulated in a number of ways, some of which will be mentioned later. But quite apart from the method of stimulating the box so that waves are contained therein and currents flow along the walls, this cavity will have a frequency that may be called "resonant" for the oscillation described. In other words, there is a frequency for which there will be very high electric-field amplitude between the plates of the capacitance, at least in the center of the

box, because of the charges existing on the top and bottom, and a very high current flow up and down the sides of the box. At some time the energy will be stored in electric field. At a later time it will be in magnetic field. Just as in the simple LC resonant circuit, the wall currents and the top-to-bottom voltage will be out of phase by 90 deg.

When the word "circuit" is mentioned to a centimeter-wave engineer, the picture that comes to his mind is not likely to be the same as that which comes to the mind of an engineer working in the radiobroadcast or power range. The former's concept of circuit is associated with the properties of enclosed regions. The great importance of the resonant cavity in the ultra-high-frequency region, because this form is the commonest of the circuits used there, makes it necessary that the engineer have in mind the various resonant modes, the various patterns of electric field and magnetic field, and current and charge distribution that can exist in those cavities. Hence some of the thoughts of a microwave engineer on electric circuits are similar to the thoughts of an acoustic engineer about cavities containing sound waves. This statement does not imply that centimeter-wave electricity is very closely related to sound waves. Because for both phenomena the physical sizes of the systems with which the engineer works are of the order of the wavelength, it is not uncommon to find very similar phenomena in the case of both centimeter waves and acoustic waves.

We arrived at this picture of charge and current distribution inside a box under resonant conditions by extrapolating from the coil and condenser, but we could just as well study the cavity by thinking of the space inside the box as one in which centimeter waves are propagating. The waves are not able to leave; therefore, if they are to exist inside, they must bounce back and forth between walls. The consideration of the resonant cavity as a region where waves are propagating instead of as a circuit will be left for a later chapter. But it may be noted here

that, in order that the wave may fit into the box, the electric-field strength, which may be high near the center, must decrease so that at the sides (where it is in effect shorted by the conducting wall) it will be very small. It is necessary, in other words, to have half a wavelength or some multiple of half a wavelength across the face of the box (Fig. 49). If the frequency is not high enough so that the wavelength is low enough to fit into the cavity, then the cavity cannot resonate. It may be excited at this lower frequency, but it will be far off tune, and the response will be low.

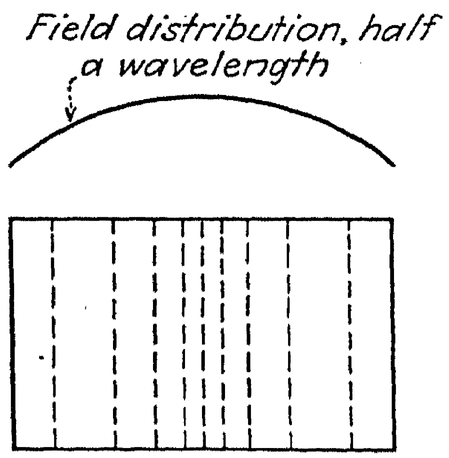

FIG. 49.—If the frequency is not high enough (wavelength low enough) for half a wavelength to fit into the cavity, it will not resonate.

Now let us consolidate our gains. We have called a cavity, the inside of a conducting box, an "ultra-high-frequency circuit." This sounds reasonable for the moment because we remember that the resonant cavity grew from a simple one-turn-inductance–one-condenser parallel circuit. Yet this is not completely convincing until we see how it is used, which we shall do shortly. First, let us remember that we came upon this type of design because of a desire to eliminate radiation. That we have indeed eliminated it is clear if we recall that if the walls of the box are perfect conductors, because of skin effect, no electromagnetic field or electrical effects or energy will be able to get from the inside of the box to the outside. (Except, of course, a static magnetic field.)

Current Distribution.—Let us continue with the simple, perfectly conducting box. If the top and bottom of the box are charged oppositely at some instant, the charges will, of course, attempt to come together to neutralize themselves (Fig. 50). This will set up along the sides of the box currents that will flow on the inside, *i.e.*, if the source is on the inside. The discussion of Chap. VIII

indicated, at least in the case of the simple one-loop circuit, that when a circuit does not radiate energy, then the current distribution cannot be uniform all the way around the circuit. This is a thesis that seems to fit this case. Here we have a problem in which we are certain from fundamental skin-effect considerations that there can be no radiation. Therefore, no matter how the stimulation of the cavity may take place, it is fairly certain that there will be no uniformity to the distribution of current. Such a picture fits the fact that the current flow is mainly on the sides of the box. The top and bottom have charge distributions, so the current flow on different parts of the top and bottom must be different. There will, for example, be essentially no current flow toward the center of the top or bottom (Fig. 50). The function of the current is, from one point of view, to bring the charge up to the top or to the bottom. On the walls of the box there will be no charge, and consequently the current will be the same all the way up the walls.

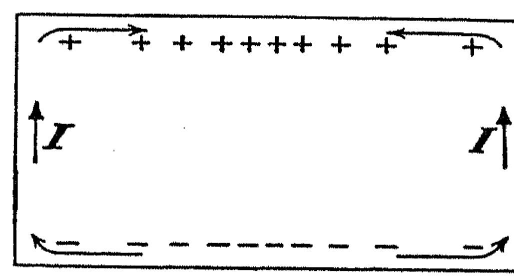

FIG. 50.—The current flow is mainly on the walls of the box.

Electric-field and Charge Distributions.—Now that the question of current distribution has been brought up, all the distributions—currents, electric field, charges, and magnetic field—should be listed. First let us take the electric field. It exists between the top and bottom plates. It is obviously not uniformly distributed. It will be strongest at the center; it must fall off as it approaches the sides, as Fig. 49 has already indicated, because, with a perfect conductor assumed, the electric field, if it existed parallel to the wall, would cause infinite current flow there. When the electric-field distribution is known, the charge distri-

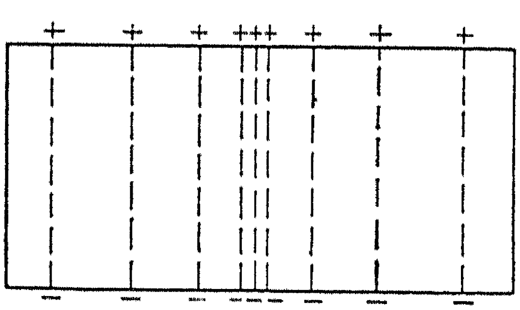
FIG. 51.—Electric flux lines end on charges on the top and bottom plates.

bution follows immediately. Where the electric field ends on conductors, there must be a distribution of electric charge whose strength is, of course, proportional to the strength of the electric field (Fig. 51).

Magnetic-field Distribution.—Now for the magnetic-field distribution—this is not so easy. We must recall that there is no a.c. effect whatever on the outside of the cavity because it is assumed that the sources are all inside. Yet, since there is current on the walls, we might at first fear that a magnetic field might exist on the outside; for the wall current would appear to be perfectly capable of giving rise to such a field. If a circular loop is drawn around the outside of the cavity (Fig. 52), there is current linking it. But displacement current must not be overlooked. We can think of the current as traveling up the walls of the box from the bottom to the top plate and returning as displacement current to the bottom plate. Thus, if a loop is drawn outside the cavity as shown in Fig. 52, and if it is claimed that there ought to be some magnetic field around the loop since it is linking current, it can be answered that both a going current and a return current of equal amounts are being linked. No value other than zero can be found for the magnetic field. The magnetic-field distribution inside the box follows at once. It simply links the displacement current flow as shown in Fig. 53.

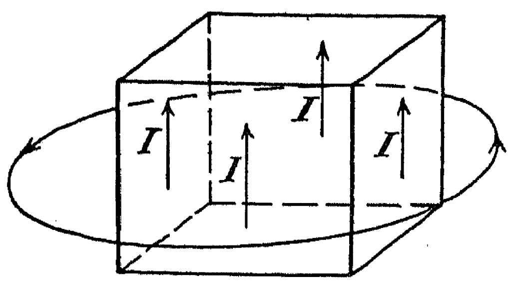

FIG. 52.—The encircling loop links a total cavity current of zero.

Now that it has been seen how displacement currents, electric and magnetic fields, conduction currents, and charges work together to yield a consistent picture of the operation of one resonant cavity, the breadth of the entire subject should begin to make itself apparent. Obviously, the example treated is only the beginning. Quite a number of different distributions and hence different

eventual circuit characteristics are possible in any space enclosed by a conductor. Figure 54 shows cross sections of rectangular parallelepiped cavities in which two other modes of distribution are stimulated. The dotted lines

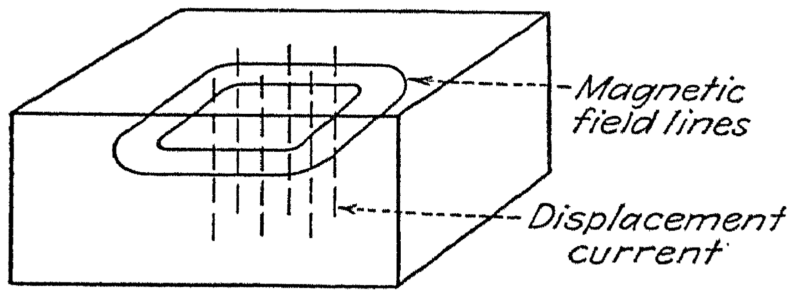

FIG. 53.

represent electric-field flux lines connecting the positive and negative charges pictured at one instant of time. These ways in which electric field may distribute itself inside cavity resonators are as important to the centimeter-

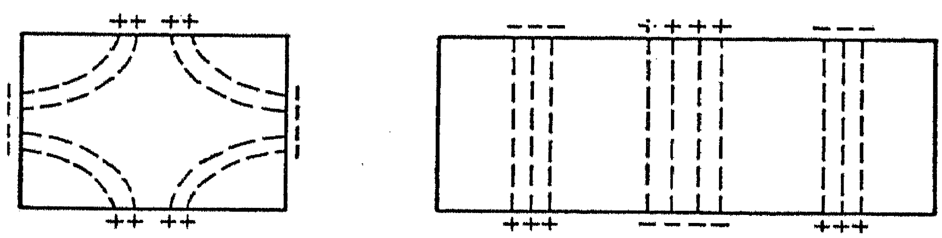

FIG. 54.

wave engineer as Y and delta connections are in the power field, and tuned transformers in the radiobroadcast receiver field.

Notice how we can obtain modes in which there are several nodes and loops of charge distribution or voltage distribution as we investigate the pattern from the center outward.

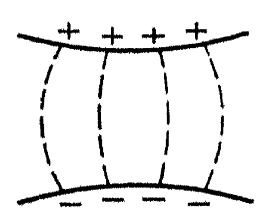

FIG. 55.

It is easy to fashion many other patterns of electric fields and magnetic fields and the resultant cavity boundaries that will enclose them. Where electric-field lines end, we will put conductor surfaces to hold the charges (Fig. 55). Then we run conduction current up along the walls of the conductors to feed the charges (Fig. 56). This closes the system. We now investigate the magnetic

field by looking into the linkages of the total current displacement current plus conduction current (Fig. 57) As an example, suppose we desert the straight electric-field lines of most of the preceding drawings and sketch a few curved lines of electric field as a beginning. This is indicated in Fig. 58. Next we shape conductors on which these lines can end, the lines coming in perpendicular to the surface; finally, we close the conductors to make

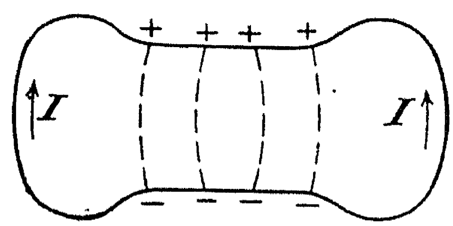

FIG. 56.

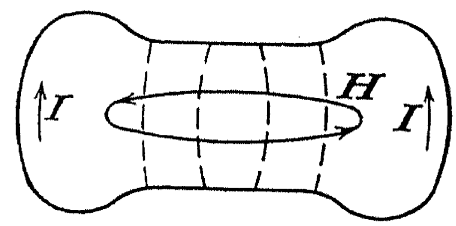

FIG. 57.

possible the conduction current to carry the charges up and down between the two plates on which electric-field lines end. Again, the dotted lines show the electric-field distribution at some instant of time, and the solid lines indicate magnetic field. Figure 58 can be interpreted as showing, crudely, the cross-sectional distribution in the axial plane of a spherical cavity (Fig. 59) or in the trans-

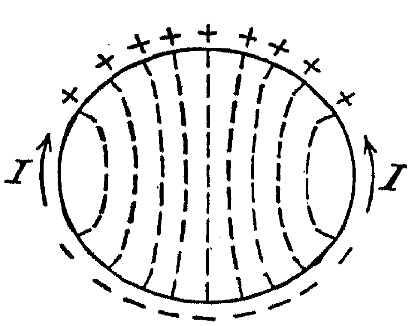

FIG. 58.

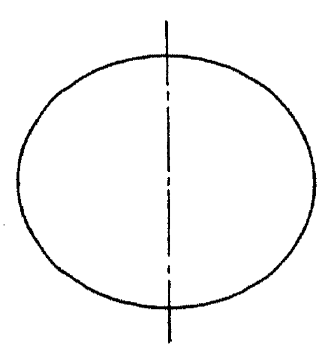

FIG. 59.

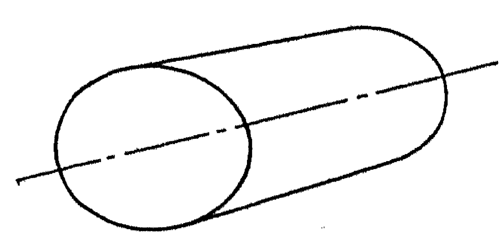

FIG. 60.

verse plane of a cylindrical box cavity (Fig. 60). In the latter case, note that it is the sides of the cylinder that are charged against one another; in the former, it is the poles of the sphere. The electric field will be strongest in the central region of the cylinder. It will have to be very small (zero, if the conductors are perfect) at the two ends of the cylinder, where the electric field will be shorted by the conducting end planes.

We can repeat the above process, with the electric-field lines bending the other way and end with the biconical cavity of Fig. 61.

Of course, it is not necessary to have the electric-field lines ending on conductors. In the preceding chapter the general field laws were discussed. One of these allows the electric lines to close on themselves, linking the magnetic-field lines, in direct analogy with the magnetic lines of the previous illustrations that link the electric-field lines. For such a situation to exist in a cavity, there must be, as usual, a proper distribution of both the magnetic- and the electric-field lines in order that the changing magnetic field may be able to induce the changing electric field around the linking path, and vice versa.

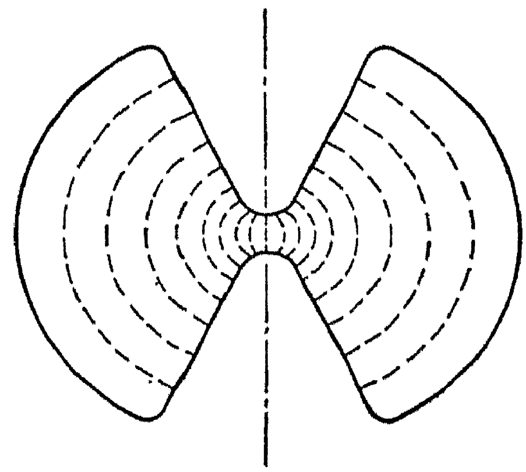

FIG. 61.

The reader may legitimately sense at this point that it is easy to oversimplify resonant cavities if they are thought

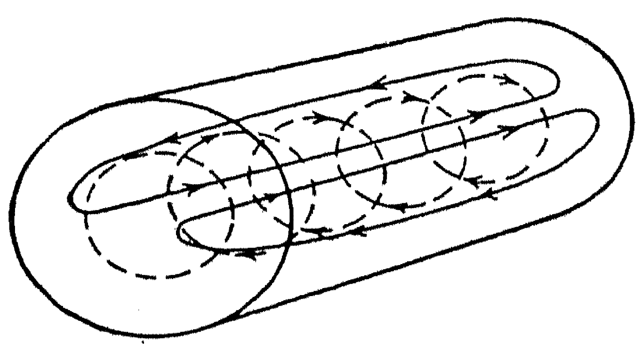

FIG. 62.—An oscillation mode in which the electric-field lines (dotted) also form closed loops.

of purely as an extension, to prevent radiation, from a simple *LC* circuit. Now that the possibility of having electric-field lines closing in loops—not starting from one place on a charged conductor and ending on another—has been introduced, it becomes apparent that the complete resonant-cavity picture is a broader one than the *LC* circuit extrapolation would indicate. For example, in Fig. 62 a cylindrical cavity is shown operating in a mode

in which the electric-field lines are circles, concentric with the circular conducting boundary, strong in the center of the cylinder, weak or essentially zero at the two ends. The magnetic-field lines travel axially and radially so as to link the electric field.

This example certainly does not seem at all analogous to the condenser-plate situation. Indeed, it bears some resemblance to a shielded solenoid (Fig. 63) because the displacement current is as effective as the solenoid's conduction current in producing magnetic field. Even this similarity is not worth pressing. It is not necessary

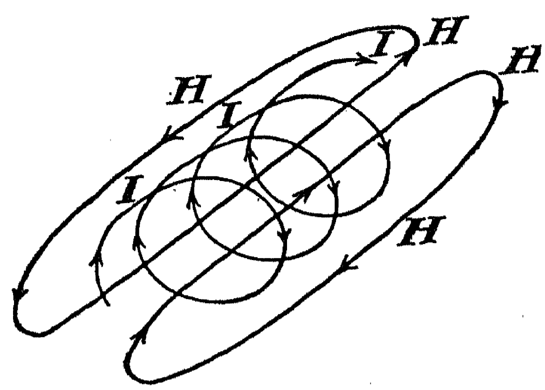

FIG. 63.—Two of the magnetic field lines H linking a helical current I.

to hold to analogy to understand these pictures. It is simpler to give free license to the view that the fields may mutually stimulate one another. After all, with a cavity having a perfectly conducting boundary, it is easy to imagine that somehow or other one of the fields, say the magnetic field, gets started. As it starts falling off in strength, it will induce an electric field because the induced e.m.f. around the path indicated by the dotted lines of Fig. 62 will be proportional to the rate of change of the linking magnetic field. Thus the electric field will rise in strength as the magnetic field drops off. As the magnetic field passes through zero and changes direction, the electric field will reach its maximum and fall off. Now it can be correctly stated that the changing of the electric flux, or displacement currents, gives rise to a changing magnetic field that will in turn support the original mag-

netic field carrying it through zero to a reversed peak. Thus the whole system stays in resonance, the energy oscillating back and forth from electric to magnetic field and choosing a frequency for which the electromagnetic-field distribution will be able to fit properly and comfortably inside the particular size and shape of cavity that has been chosen.

This discussion of the way in which an electromagnetic field can maintain itself if there are no losses is an excellent way to introduce the subject of wave propagation in a more detailed way than has been done so far. As a matter of fact, this is the approach that we shall use in the next chapter. It is easily granted that inside a perfectly con-

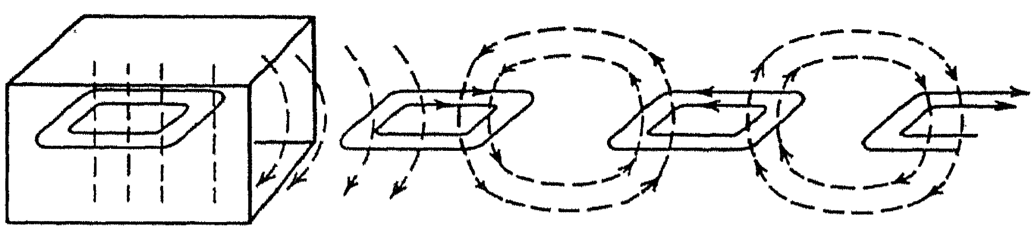

FIG. 64.—With the cavity wall removed, a coupling would exist to outside space because of the mutual induction of electric and magnetic fields.

ducting box the electric and magnetic fields will keep themselves up in just the same way as the magnetic stored energy and electric stored energy in a parallel L and C. If the box were opened, if one end plate were removed as in Fig. 64, the energy might be expected to move off as waves through space, the electric and magnetic fields inducing one another forward.

An interesting thing about these self-enclosed circuits is that the ratio of stored energy to energy dissipated is unusually high for ordinary conductors compared with the lumped circuit used at low frequency. Not only has radiation been diminished to negligible amounts, but this type of circuit construction is better in efficiency of energy storage than a typical low-frequency resonant circuit where radiation never was a factor. This is perhaps not a thing to prove rigorously here, but it can be made to sound reasonable. Such current flow as there is, is distributed

over a relatively large area. This tends to cut down the ohmic losses on the imperfect conductor's surface.

Of course, the type of field distribution indicated in Fig. 62 simply could not happen in statics, but perhaps it is foolish to bring up such a point. It certainly can happen at any low frequency if only the dimensions of the box are made large enough.

If the microwave engineer's notion of a circuit consists in a large measure of pictures of oscillation modes inside a closed region, does the operation of these circuits in practical systems involve material differences, compared with the use of circuits at the lower frequencies?

It is probably safe to say that, in the utilization of these circuits, the functions of these cavities are always analogous

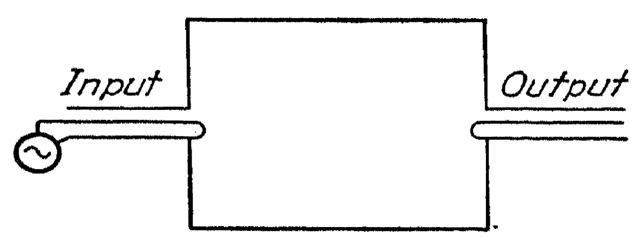

Fig. 65.—A simple filter using a cavity resonator.

to the low-frequency circuit. For example, Fig. 65 shows a simple filter. Energy comes in by the opening shown with a loop to excite the magnetic field. If the energy arrives at a frequency that resonates with the box, very high current density and charges will appear on the cavity's inner faces. Obviously, the construction of the input loop and its location in the cavity has a great deal to do with the type of mode that is stimulated. If the frequency is not very nearly one for which one of these modes would like to exist, then very little energy will be transferred from the loop to the cavity, and the resonant cavity will have exceedingly low currents and fields. Consequently, very little magnetic field will link the output loop.

Figure 66 shows two electrodes of a vacuum tube, two grids through which an electron beam will pass. The

current in that beam is assumed for this illustration to contain variations in density. As each lump of space charge passes by, it induces current in the external circuit connecting the two grids according to the principles discussed in Chap. V. Also, for reasons that stem from

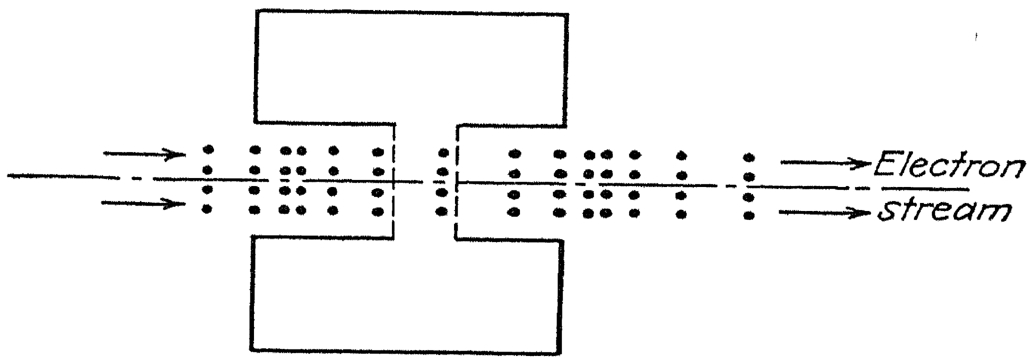

FIG. 66.—The axially symmetric cavity resonator is connected across the two grids.

electron transit time discussed in Chaps. IV and V, the two grids will be rather close together. Thus the choice of shape of resonant cavities that are used with microwave electronic tubes is accounted for. One suitable cavity form is indicated in Fig. 66.

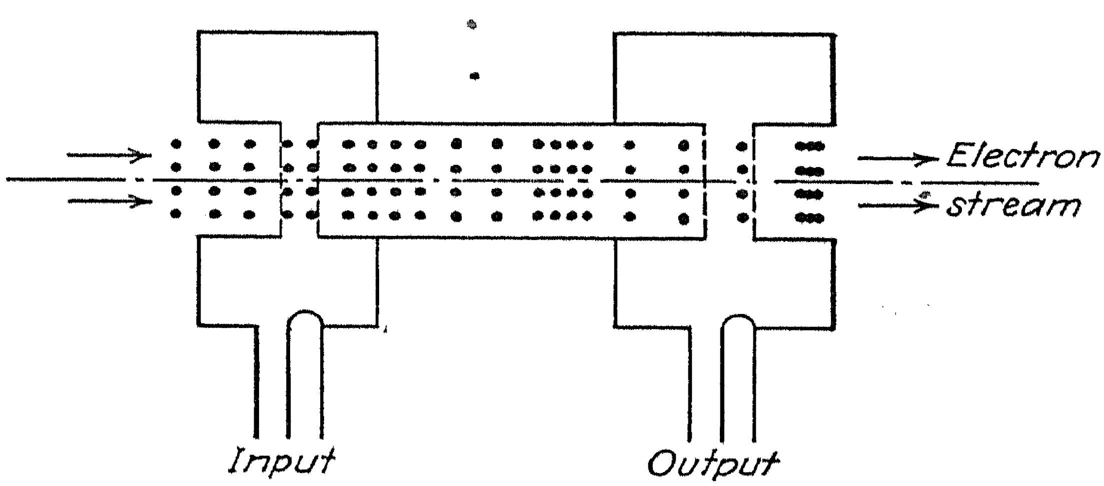

FIG. 67.—A two-gap velocity modulation amplifier tube using two resonant cavities.

Figure 67 shows a two-cavity velocity-modulation amplifier, one cavity being excited by input power with the excitation and cavity form so selected as to yield highest signal voltage across the input gap. This signal voltage acts on the electron stream (Chap. IV), which, as it moves

from input gap to output gap, develops density variations. As these variations pass the output cavity's gap, they induce current in, *i.e.*, excite, that cavity. Then this output-cavity power may be drained off by a suitable opening from the cavity to an outgoing transmission line.

CHAPTER X

All Conducting and Dielectric Boundaries Are Wave Guides

In each of the previous chapters, an attempt has been made to introduce at least one important concept applicable to the higher radio frequencies. It is time now to draw a major conclusion. Perhaps it is premature to call it a conclusion; it is, at any rate, a concept and will be the subject of this chapter. The approach adds nothing new;

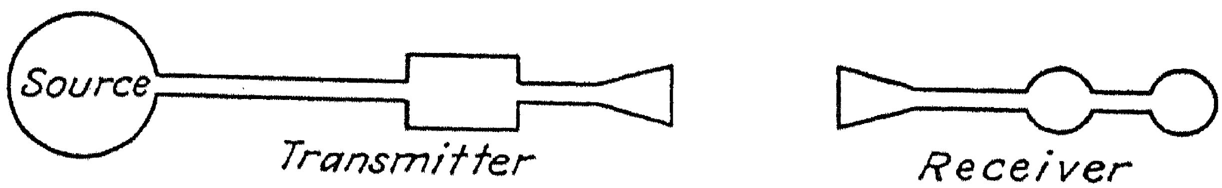

FIG. 68.—No effective break is allowed in the shielding of a microwave system until intended radiation takes place.

it rather attempts to correlate all that has gone before and to lay the groundwork for a better approach to what is yet to come. The theme for this chapter is: All microwave systems, all electrical systems, at any frequency, are *wave guides*. At the lower frequencies, this concept is perhaps not the most valuable approach, except, of course, for antennas and transmission lines. At least, it is not a necessary approach for all electrical systems, and it may be a cumbersome way of looking at some parts of the system, such as the simple circuit of low frequency.

It is well to approach the development of this theme by noting first what is meant by the term "electrical systems" at ultra-high frequencies. Figure 68 shows diagrammatically a general system that is typical of microwave setups. Power flows out of the source. It would spread as waves in the space surrounding the source except for the fact that the apparatus is made self-shielding; *i.e.*, the source is

some sort of electronic generator whose oscillating circuits are resonant cavities. Energy is abstracted from these by connections to hollow pipes or coaxial cylinders or some equivalent wave-guiding system, care being taken to see that skin effect maintains the microwaves on the inside of these guides. There is no actual break in the skin that covers the microwaves, *i.e.*, not until it is actually intended that the waves be radiated. Then, certainly, a hole appears in the over-all enclosure, and this opening is very carefully designed to send the waves out properly or efficiently and in the intended direction. A similar situation occurs, in reverse, at the receiver end.

At one or more points between the tubes and the antennas there may be resonant regions that will serve the usual purposes of selecting frequencies, coupling the longer guides together, etc. What happens at each step depends on the boundaries of the space. Sometimes the energy is contained in a relatively small region of space. Then the electrical system is usually called a "circuit." Sometimes the energy that leaves the source is guided by conductors and dielectrics to some fairly distant place in space where it is reflected or absorbed. Then the system is usually spoken of as a "transmission line" or "wave guide." Sometimes the boundaries around the source are especially selected to entice the energy out of the source and start it off into the surrounding space so that it will be transmitted through that space without benefit of further guiding. Under those circumstances, the electrical system is spoken of a "radiator" or an "antenna."

But we have seen that, as a general rule, all circuits may radiate unless they are rather special, say completely enclosed by perfect conductors. Also, we have seen that every simple conducting system that might ordinarily be considered as a circuit is in a very definite and proper sense a transmission line, since energy is guided from the source to some terminal, even though that terminal may simply be the other end of the circuit where most of the energy is

reflected. Everything then comes down to a matter of the order of magnitude of the effect—the actual numerical ratios between the radiated power and that which is held by the conductors in their vicinity, or between the radiated power and the power that is taken from the source and deposited in a load with the benefit of guides for the entire distance.

It is safe to use the wave idea in correlating all three systems because it has been seen that it is completely legitimate to consider that all electrical energy resides in the dielectric bounded by the conductors; no high-frequency waves are able in a practical way to penetrate the conducting surfaces.

Of course, this picture of all electric phenomena as waves has been built up throughout the text gradually. Also wave and field notions have been very carefully correlated with the conventional voltage and current ideas, because no concept that comes in at ultra-high frequency can do so suddenly. If we continue to use correct concepts throughout, then we must be able to see always how the low- and the high-frequency concepts fit together. We shall continue to do that throughout the following chapters, in which we have yet to add the discussion of transmission lines or their equivalent and antennas and to correlate both of these with circuits, especially the resonant cavities studied in the previous chapter.

Importance of Boundary Conditions.—The importance of boundary conditions around a region of space that has been caused to put up with very-high-frequency electrical phenomena can best be understood by studying some specific examples. First, let us say something more about the relatively easy situation of a completely enclosed system. Figure 69 shows a coaxial transmission line consisting of two presumably perfect conductors. A source of voltage is applied between the two conductors. As a result, electromagnetic effects will be transmitted down the transmission line (or, to use the term preferred in

this chapter, the "guide"). Since the source is on the inside of the space bounded by the two conductors, no effects will be at all apparent on the outside. Thus the only fact that will be of concern is that waves will be transmitted down the region of dielectric that is bounded by the two cylinders. Now, if, as Fig. 70 shows, the coaxial-line system is twisted in a most confused way, the energy will still be restrained by the action of the perfectly conduct-

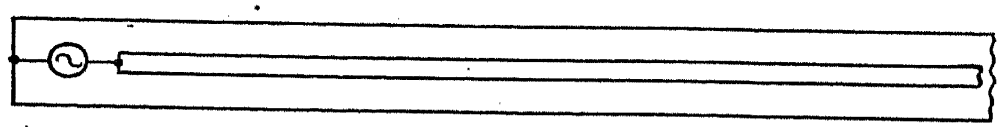

FIG. 69.—A coaxial line with the source on the inside guides power in the inter-cylinder space.

ing boundary to stay within the guide. Notice that there is no comment on the efficiency of this system; for instance, nothing is being said about how much energy might be enticed around the curves. It may be that when a wave arrives at the curve, only an inappreciable amount of the wave energy may continue around the curve with the rest of it being reflected back to the source. On the other hand, perhaps almost all of it will be persuaded to go

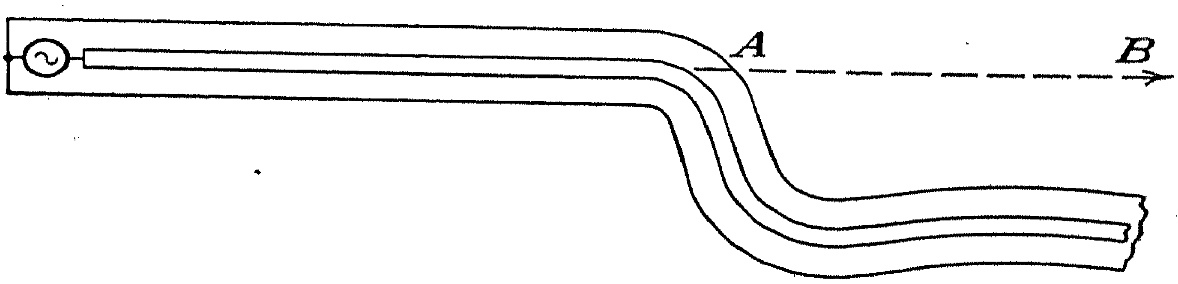

FIG. 70.—The twisted coaxial line still holds the waves inside.

around the curve. This all depends upon the details of the particular choice of dimensions and frequency, the amount of the curvature, and what there is at the other end of the curve. The point is that the energy is restrained and guided to follow the form of the boundary. In no case are microwaves found outside the region bounded by the two conductors. No waves continue in the original direction AB (Fig. 70), ignoring the bend in the line.

The conductors insist that the electric fields be perpendicular to their surfaces. (Again, if there were any electric field tangent to the perfect conductor, an infinite current would be caused to flow in it since it has no resistivity.) If, as pictured in Fig. 71, the electric field attempted to go on through space instead of taking the curve, there

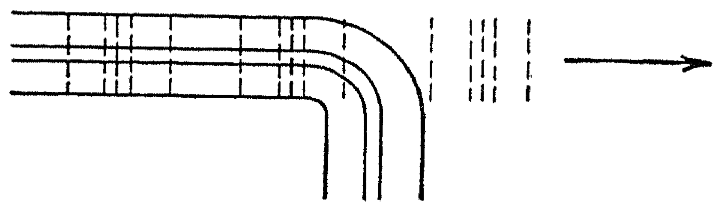

FIG. 71.

would immediately be a component of that electric field parallel to the conductor. This is complete proof that no such secession by the electromagnetic waves is possible, but it is the same as citing the skin effect again.

A final example of how a perfectly conducting boundary guides electromagnetic waves is found in the resonant cavity already considered in Chap. IX. The cavity shown in Fig. 72 is axially symmetric and contains a small gap on the axis across which a high-frequency voltage is placed. Electric-field lines that result as part of the manifestation of the excitation will be kept perpendicular to the conductors; no effect will be observed at all on the outside of the cavity, since it is assumed that no additional source is there. The wave-guiding way of looking at the cavity action is approximately as follows: The waves will travel radially outward from the central source, guided by the conical surface AA. At the spherical boundary BB the waves are reflected. This is a way of recognizing that the electric field must be zero tangential to BB. The reflected waves and the incoming or incident waves will

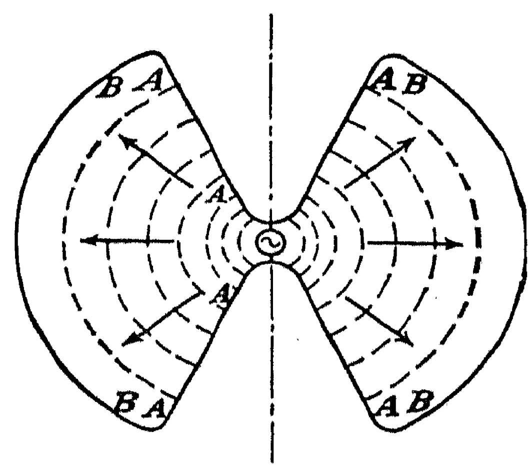

FIG. 72.—The waves travel radially outward, guided by the cones, and are reflected by the sphere.

be of such phases relative to one another that the total electric field from both waves (1) will always be zero on *BB* and (2) will always satisfy the generator condition at the source.

This is a steady-state situation. Waves are present in the cavity at all times, and they can be divided into two classes, those guided outward to the end of the cavity and those guided backward after reflection. The two waves add together to give the current, charge, and magnetic- and electric-field distributions that were agreed in Chap. IX to exist. In that previous discussion, the distribution was reasoned out directly from the two fundamental laws that told how electric and magnetic effects induce and

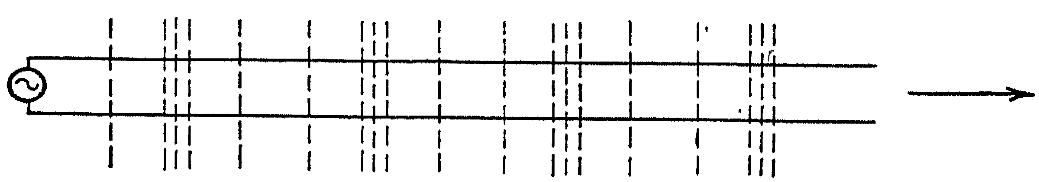

FIG. 73.—An open two-wire line also guides waves.

support one another. Now the same problem is looked at as one of guided waves, waves that are the general result of the idea that the electric and magnetic effects induce and support one another. The same answer is certain to be obtained in each case.

Open-wire Guide.—So far only completely enclosed guiding systems have been considered. In such situations the microwaves simply could not escape. Their only choice was between following along the inner boundaries and being reflected back to the source. A more complex guiding system is the open-wire guide or transmission line pictured in Fig. 73. Here again is seen a source of voltage that results in electric field between wires. There are, of course, also magnetic field and currents and charges, but the electric field will be noticed particularly. It comes in perpendicular to the conductor. Thus, if the conductors are twisted around (Fig. 74), there is the same guiding action noted before. The electric field near the conductors

must turn around with the conductors so as to remain perpendicular to them at all times. The waves, guided by the line in the original direction AB, will be pulled around the corner and thenceforth guided to travel in the direction BC; *i.e.*, some or most of the wave energy may make the turn.

Unlike the previously considered self-shielding systems, the open-wire line allows the fields to be spread over all the surrounding space. As long as no twists or discontinuities exist in such a system, if there are only two parallel perfect conductors, then the wave energy will be confined to

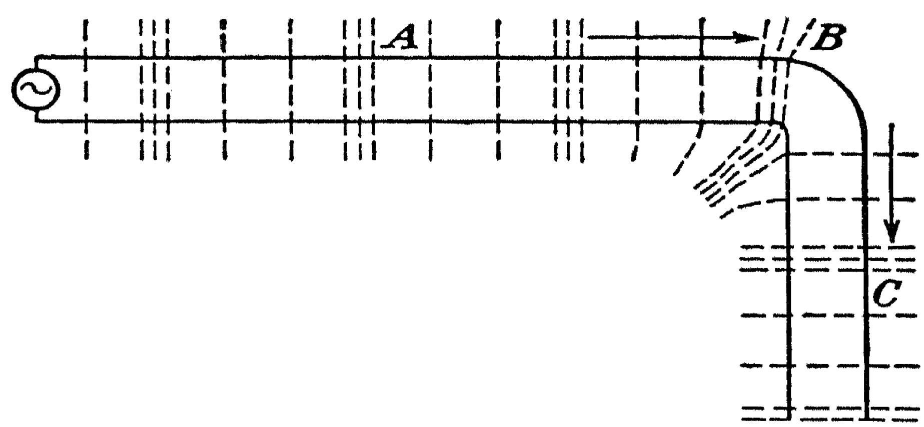

FIG. 74.—Some of the power is guided around the turn by the open line.

moving along the system. Despite the fact that the fields go out to infinity, the direction of power flow will be in the direction of the guide. Under these circumstances, the guiding idea is sufficient, and no loss of energy takes place in any unguided waves that simply decide to travel away in some other direction. But such a condition is an impossible one. It would certainly be impossible to fashion the source in such a way as to start the fields out properly. Bends and end effects at the load end will also cause distortions.

At each twist or discontinuity, the electric-field distribution near the conductor will be closely dictated by the boundary surface. But since the electric field spreads out to infinity, the more distant energy will not notice the wires appreciably. If this matter could be investigated further, it would be seen that when there is a discontinuity

in the open-wire guide, the waves do indeed continue in part along their original direction. The electric and magnetic fields have been carrying each other along in the direction of the line. This is true even for the fields far out from the wires. When the twist occurs, the conditions of mutual induction toward a motion in the original direction still exist. A little farther on, past the turn, the effect of the turning of the neighboring waves—those closer to the wires—will be felt by the more distant fields. Thus the distribution of the stubborn distant waves will be altered. But they will still have a component of flow in the original direction. After all, only the field on the wire is obliged to come in perpendicular to it.

More accurately, the twist will cause waves to disperse in all directions. Some will make the turn; some will be reflected back to the source. The nonconforming waves will go partially in every other direction. Such waves are lost to the line forever. Of course, if they should strike some other distant boundary, some of the resulting reflection might be recovered by the line. The line has become an antenna. Waves are radiated from it as well as guided by it.

Obviously, the waves that travel in the original or nonguide direction and that are spoken of as "radiated" are given that title as a matter of definition. They are the unguided waves. Whether they are useless or useful depends upon what result is sought. Careful design of the leaky guide would bring about a good transfer from guided to unguided waves and would direct the radiated waves where desired. Then a practical antenna would result.

CHAPTER XI

Most Engineering Concepts of Transmission Lines Are Valid for Microwaves

The previous chapter showed how all microwave systems are partly wave guides. Every conceivable component of a microwave system not completely enclosed seems to be part circuit, part transmission line, and part antenna. Despite this possible and even probable hybrid nature of things at the higher frequencies, it is still extremely useful sometimes to separate systems into these different classes, provided that such a classification is not allowed to make the separated groups appear unrelated. One of the most useful classes at any frequency is the common two-conductor transmission line.

For present purposes, a system will be called a "uniform transmission line" if it has two parallel constant cross-sectional conductors whose lengths are substantially greater than their spacings and at least comparable with the wavelength. Uniform lines are used at ordinary power and communication frequencies to transmit power and intelligence. They are used for the same purposes in the microwave region, except that here the single hollow pipe or common wave guide, studied in the next chapter, offers competition. But transmission-line concepts are much more important at the extremely high frequencies than these statements have so far indicated. When the wavelength is best given in centimeters, the distance between two component parts of a piece of apparatus is almost certain to be a substantial part of a wavelength. At microwave frequencies any little interconnecting lead is a candidate for the transmission-line class.

In addition to this rather obvious importance of transmission-line theory at ultra-high frequencies, there is another reason why a separate chapter should be devoted to a consideration of transmission-line concepts. There is a conventional, seasoned approach to transmission lines which engineers have found useful and satisfactory for the low and medium frequencies. Much of this approach is completely practical and legitimate even for microwaves, for reasons that help greatly in the understanding of concepts of fields and waves and circuits.

Transmission lines are to be considered in the next few pages, with the accent, of course, on concepts. Some things will be proved, and others will be stated with only

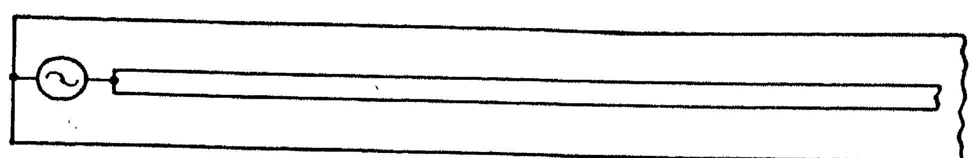

FIG. 75.—The favorite line at extremely high frequencies is the coaxial line.

a few accompanying comments to make the points seem reasonable. Still other discussions will deal merely with tying in the transmission-line class with other classes of systems, in order not to lose sight of the fact that the division into classes is not necessary.

Discussion of the first concept applicable to lines at the highest frequency may seem almost entirely repetitious. Yet its importance requires at least a summary statement at the outset. The self-shielding, radiation-reducing idea that formed the basis for resonant cavities is equally applicable to transmission lines. To ensure the transfer of power without extraneous couplings to other systems or other parts of the same system, a transmission line is generally chosen that has a complete shield about the propagating fields. The favorite is the coaxial line shown in Fig. 75. If the excitation and loading of the line always take place in the space between the conductors and if the outer conductor is not broken anywhere, then there is no effect on the outside, practically speaking.

Throughout this chapter, unless special mention is made to the contrary, it should be assumed by the reader that the discussion is free from the complications of possible radiation from the line and annoying couplings from the line to its surroundings. The fields will be conceived of as bounded uniformly at each cross section by a surrounding conductor that most often, in practice, will be one of the two conductors of the line.

The next fact that deserves a place here is that transmission lines are used to replace lumped circuits at high frequencies as often as they are to transfer power from a

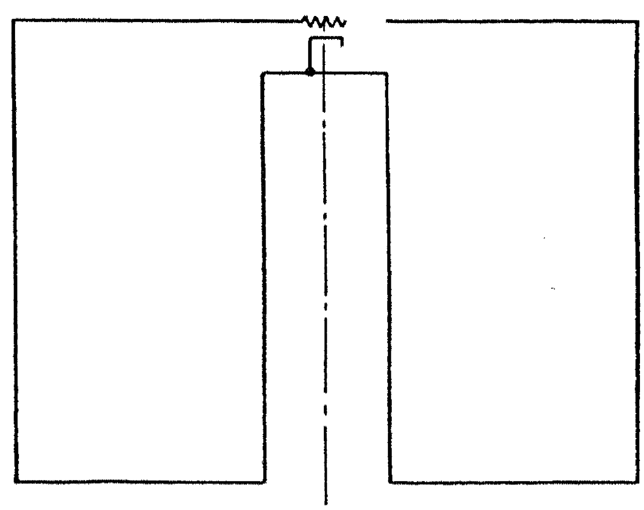

FIG. 76.—A shorted length of coaxial line connected between grid and cathode of an electron tube.

source to a load. It is common to speak of the use of sections of lines as "circuit elements." They are used as the tank circuit of an electronic tube or as transformers to convert one impedance into another. Often, when these sections are critically sensitive to frequency, they are called "resonant lines."

As one example, Fig. 76 shows a section of coaxial transmission line that is shorted at one end and connected across two terminals of an electronic tube at the other. If the line is about a quarter wavelength long at the operating frequency, it turns out that the waves reflected by the shorting plate reach the tube on their return in such phase as largely to cancel the current in the outgoing wave.

Under these circumstances, the line presents a high impedance to the tube. Since this happens over fairly narrow frequency ranges (where the line, of constant dimensions, comes out a quarter wavelength, three-quarter wavelength, etc.), the line is roughly the equivalent of an ordinary LC tank (Fig. 77) near its resonant frequency.

The reader may say that Fig. 76 could just as well have been a picture of a resonant cavity so shaped as to make it fit conveniently around the two terminals of the tube. In fact, if this thought does not seem perfectly sound to the reader, then the previous chapters have failed at least partially in their objectives. Surely, if a piece of a self-enclosing type of transmission line—and it does not even have to be uniform—is cut out of a long line and the ends are closed by conducting plates, the region enclosed by this effort is a resonant cavity. The general ideas of Chap. IX apply. So, when all is said and done, it is often a matter of taste and sophistication whether microwave engineers speak of many of their circuit equivalents as "tuned lines" or "resonant cavities."

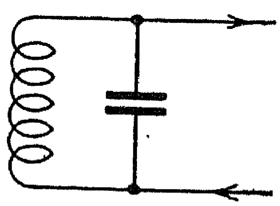

Fig. 77.

Many engineers, accustomed by years of experience to the use of transmission lines as circuit elements, believe it somewhat of an affectation to speak and think in terms of cavities. Their contention is that the hollow boxes are rarely as close to the true system as is the short-circuited line of Fig. 76. Of course, the purpose here is to see the relation of these various concepts. Each is a correct one, properly interpreted, and that is all that needs to be judged in this text.

Distributed-constant Concept.—Next to be discussed is the matter of the well-known distributed-constant concept that has served so well in designing and understanding transmission lines. Before outlining just what is to be shown, a word should be included as to the meaning of the distributed-constant idea. Transmission lines are most often analyzed by the use of Kirchhoff's circuit laws and the

notion that the line is a network of an infinite number of distributed infinitesimal series impedances running the length of the conductors, interconnected with infinitesimal parallel admittances between the conductors (Fig. 78). This attack gives practically correct answers (perfect answers for perfect conductors) but has often been suspected of basic conceptual errors and therefore considered dangerous, the answers it yields being potentially grossly in error.

It is not surprising to find that the distributed-constant approach seems to be arbitrary, particularly after a study of the general laws of electricity and magnetism. Such a study teaches caution about extending ideas that worked

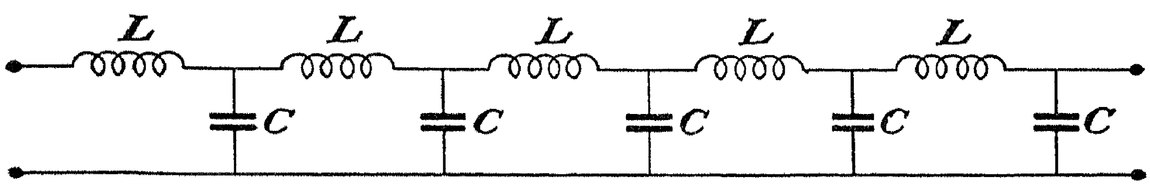

FIG. 78.—A perfect-conductor transmission line is usually thought of as a network of distributed L and C.

well at low frequency to some higher frequency. Can the concepts stand a showdown against the all-frequency laws of electromagnetics?

Yet it is just this kind of untested extension that conventional transmission-line theory appears to make. For instance, consider the impedances of Fig. 78. These are supposed to be part of an infinite network whose action represents the line, the number of elements approaching infinity as the line portion represented by each element approaches zero, the whole then yielding perfect accuracy. Take first the perfect-conductor line and deal then only with a series inductance per unit length and a parallel capacitance per unit length. How are these two distributed constants determined so that the analysis may then proceed? Consider first the series inductance. The inductance per unit length used is precisely that obtained in computing the flux linkages per ampere per unit length for the same system with direct current flowing down one

conductor and returning by the other. Also, the capacitance per unit length used in the line theory is exactly that obtained if a calculation is made for the charge on the conductors per unit length for a difference of d.c. voltage of unity between them.

From this statement of the facts it does seem as if d.c. circuit constants are being used for a problem that is anything but direct current. In fact, the results obtained from the distributed-constant approach are that the phenomena on the transmission line can be described in terms of traveling waves of current and voltage. In a typical line problem, the current and voltage will oscillate with distance at every instant. The distribution of current and voltage will be anything but constant with distance or static with time, and yet the parameters that appear to have been relied upon apply only when the current and voltage are truly constant and static.

But the distributed-constant approach is all right. It is perfect for perfect conductors, constituting the *ideal* line. Furthermore, the approach is an excellent approximate method for practical though imperfect conductors. In seeing next why the conventional transmission-line ideas are applicable and sound, two important lessons that should preface further study or appreciation of microwaves will be learned: (1) Increase in frequency does not call for a rejection of all concepts that have proved useful at the lower frequencies. (2) In particular, it is possible to interpret at least some, if not all, traveling electromagnetic fields in terms of distributed circuit constants.

Upon examination of the reason for the validity of the distributed-constant concept, it is worth pointing out that this is certainly not the first time this matter has been studied. Transmission lines have not been hiding out on towers and in cables, their true workings a mystery to engineers who were content to go along using untested concepts. That lines should be understood well is a fact that did not wait for microwaves to be appreciated. Trans-

mission lines have been subjected to thorough analysis. The ideal line has been studied by rigorous field theory—or, better said perhaps, by Maxwell's equations—and, also, by replacement of the line with an infinite network of distributed circuit constants, L and C. It has been known for years that, for an ideal line, the same answer is obtained

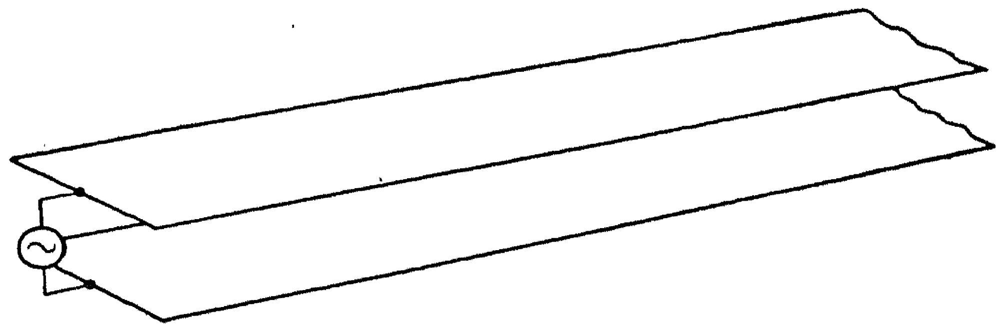

FIG. 79.—A two-conductor transmission line consisting of flat plates.

by Maxwell's equations as by circuit laws using distributed L and C identical with their static values.

Essentially, the distributed-constant approach says that the distribution of electric and magnetic fields between the conductors of a transmission line is the same over the cross section perpendicular to the line's direction as in the static, or d.c., case. Figure 79 shows a two-conductor

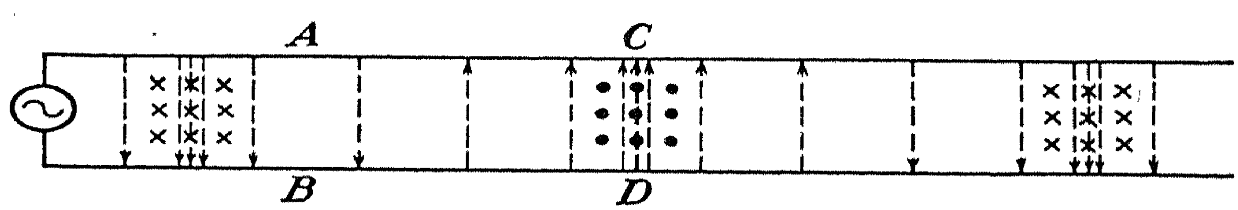

FIG. 80.—A longitudinal section of the line of Fig. 79. The electric field lines are perpendicular to the plates. The magnetic field lines (dots and crosses) are perpendicular to the electric field and parallel to the plates.

transmission line as an example to explain the previous sentence and the distributed-constant technique generally. The line of Fig. 79 is shown in a longitudinal cross section in Fig. 80. Dotted lines sketched in Fig. 80 indicate the way the electric field would be distributed as predicted by the distributed LC method. Figure 80 is a view taken at one instant and applies for the simplest situation of a single wave traveling down the line with no reflection.

The electric field is shown as entirely perpendicular to the conductors everywhere. The magnetic field is everywhere perpendicular to the electric field and to the direction of the line. It is perpendicular to the paper, in other words, and its distribution is indicated in Fig. 80 by the dots and crosses to show how it oscillates with distance along the line in step with the electric field. It must now be shown that this situation is in agreement with Maxwell's equations and that it is predictable from distributed *LC* notions as conventionally used.

The distribution of fields in Fig. 80 will be readily accepted as consistent with the two general laws of mutual field induction. Here again, except for the presence of

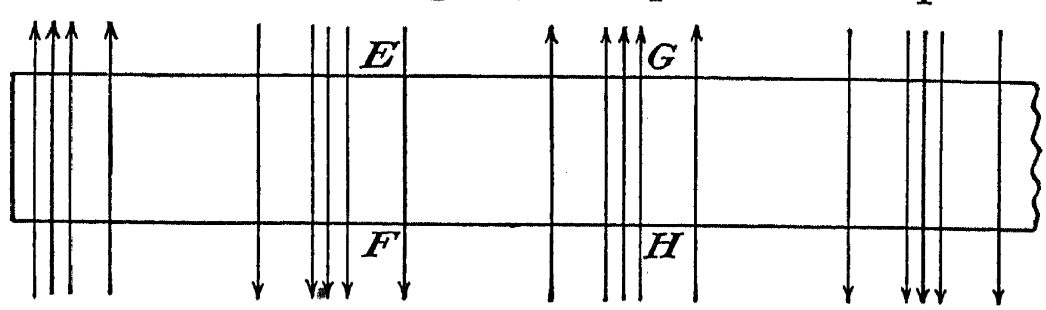

FIG. 81.—The two-plate transmission line from a plan view, showing the magnetic field lines.

the conductors, is the basic picture of the changing magnetic field's causing a changing electric field, and vice versa. The two fundamental laws can be applied and checked.

The time-changing magnetic flux passing through the loop *ABDCA* (Fig. 80) induces electric field around the loop. Hence the electric field across *AB* is different from that across *CD*. Differential elements could be added on down the line to the electric field, and in this way it would be found that, because the magnetic flux is distributed sinusoidally with distance, its oscillation in time causes an electric-field distribution that also has this oscillation characteristic in both time and distance. So far the assumed distributions of electric and magnetic field seem justified.

A continuation of the check against fundamental field theory will be simple and finally conclusive. Figure 81

shows the line from a plan view, and another loop is drawn labeled *EFHGE* around which is noted this time an m.m.f. due to the rate of change of electric flux linking it. Thus an assumption of varying electric flux down the line supports a varying magnetic flux.

This now is a complete justification of the distribution of Fig. 80. The electric-field and magnetic-flux distributions shown in that diagram are in agreement with the fundamental laws of electricity and magnetism.

Of course, this is not yet the whole story. It has been shown that it is reasonable to expect a two-conductor transmission line to be capable of guiding simple waves in the direction of the line. The term "simple" indicates that the waves are like the wave that is so common in free space and that was spoken about when the idea of retardation and wave action was brought in. These waves have their electric and magnetic fields perpendicular not only to one another but also to the direction of propagation, and everything is of like phase across a plane drawn perpendicular to the direction of propagation. The presence of the perfect conductors of the line acts to limit the extent of the field over the cross section. Also, since the electric field must come in perpendicular to the conductors, the cross-sectional shape of the conductors determines the distribution of effects over the cross section.

The next general observation that can be made about the transmission-line-type wave, a common name for this simple wave, is that the distribution of electric and magnetic field over the cross section is exactly the same as if d.c. voltage and current were applied to the line. The whole pattern throughout space is just as though a static pattern of field were moving down the line at the velocity of light.

Consider the coaxial line of Fig. 82. Alternating-current voltage is applied to the line, and as a result waves propagate down it. These waves may be described by stating that the space between lines is filled with electric and

magnetic flux. At any one instant the strength of those fields varies with distance along the line in a sinusoidal pattern. But, also, a picture of the electric-field distribution over the cross section (Fig. 83) or of the magnetic field over the cross section (Fig. 84) taken at any instant

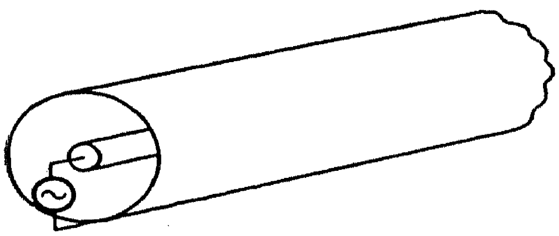

FIG. 82.

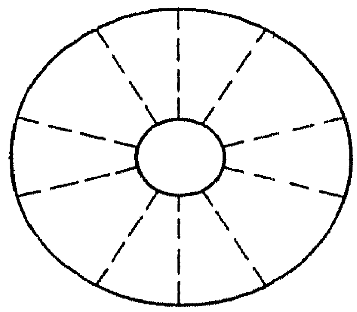

FIG. 83.

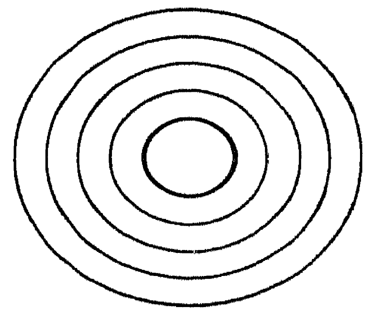

FIG. 84.

will look just like those obtained when d.c. voltage (Fig. 85) and d.c. current (Fig. 86), respectively, are applied to the line.

LC Approach.—Now not only the basic characteristics of the commonest electromagnetic-wave type have been

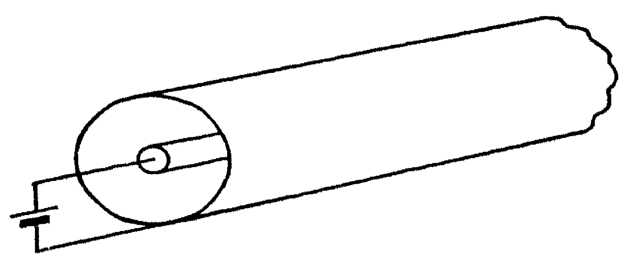

FIG. 85.

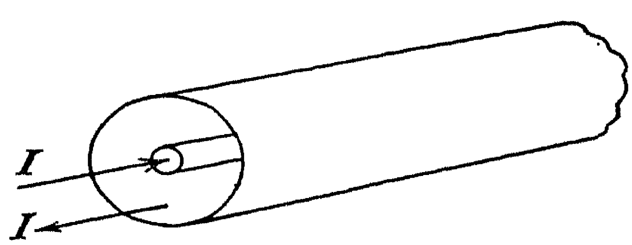

FIG. 86.

stated; but also, a link has been provided to the distributed *LC* approach to the same problem. It has been said that the *LC* method gives the same answer as does the direct study of the fields. Since the electric fields and magnetic fields have precisely the d.c. distribution over the cross

section, it can be said at once that the idea of a voltage difference to measure the integrated electric field between conductors is a completely proper one. It is a proper concept for the static case, certainly. Therefore, it is no different for the case of these particular traveling waves, which have the same cross-sectional distribution, as long as voltage differences between points on the same cross-sectional plane are spoken of.

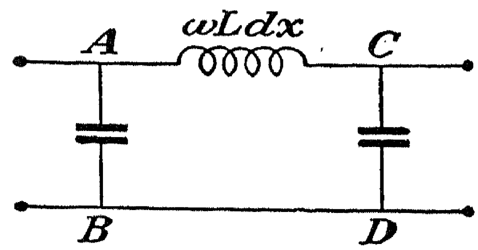

FIG. 87.

Figure 87 pictures one step of the LC approach. An infinitesimal section of line dx is shown and, according to the ordinary analysis, the voltage AB is held to be different from that across CD because of the voltage drop through the series reactance $\omega L\, dx$. But what is this L, the distributed inductance per unit length? It is only a statement of the magnetic-flux linkages per ampere of current in the line. Since the magnetic flux is entirely transverse, since it has the same distribution over the cross section as for direct currents in the line, then certainly the L may be taken from d.c. formulas, and it will still be correct. Finally, the statement that the voltage drop is current times $\omega L\, dx$ is only another way of expressing the effect of time-changing magnetic fields in inducing electric fields. To complete the LC analysis, it is now noted (Fig. 88) that the difference in current flow from one point of the line to the next point is the drain-off of parallel current flow to the capacitance between lines. But this capacitor charging current is exactly the displacement current in the changing electric flux between lines. Since the electric flux of the wave over the cross section is distributed precisely as in the d.c. case, the static C, capacitance per unit length, is a correct value to use for calculation of the charging current. Finally, since current in the line is a measure of magnetic-field strength there, it is seen that

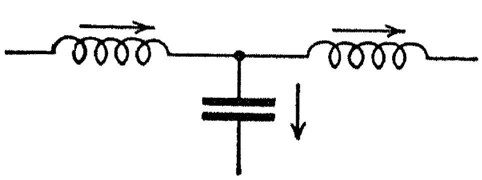

FIG. 88.

this second step in the LC analysis is just another way of saying that the time varying of electric flux induces a magnetomotive force. There is varying of current (magnetic field) along the line because of the distributed capacitance's charging current drain (time-varying electric flux).

It is apparent now why it is conceptually correct, and for the ideal line without error or approximation, to use static L's and C's to describe traveling waves of whatever high frequency. They need not be called static L's and C's. They are high-frequency L's and C's whose magnitudes turn out to be identical with the static ones.

It should also be apparent that it is unnecessary to drop concepts or results just because they originated at a frequency lower than the one that is of immediate interest. The distributed LC approach is good for any frequency—at least for the transmission-line-type wave. Of course, there are other waves; some will be discussed in the next chapters.

When the conductors are not perfect, the distributed LC attack is not precisely correct. For practical lines, however, the use of some series resistance added to the L and some shunt conductance added to the C, concepts retained, is a sufficiently exact method. The inaccuracy is scarcely large enough to justify the revision of any concepts.

CHAPTER XII

Hollow Pipes Are Practical Guides for Microwaves

One of the most attractive, and perhaps for many one of the most mysterious, types of electrical systems at extremely high frequencies is the single-pipe wave guide. With a single conductor, wave energy is guided, the waves passing down the inside, not the outside, of the guide. Here, if transmission is thought of exclusively as a matter of voltage between lines, it may be difficult to explain what goes on. With some knowledge of field theory, in addition to conventional circuits, the analysis and study of hollow-tube guides will prove straightforward. So far an attempt

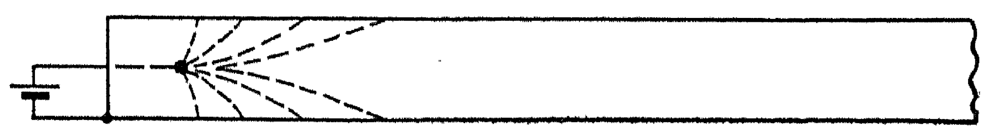

FIG. 89.—A static electric field cannot exist for a very great distance inside a hollow conducting pipe.

has been made to consolidate the basic field and wave concepts with properly used voltage and current concepts. This will be continued in describing the hollow cylinder, or common wave guide, although the whole problem could be studied by field waves alone. The inside surface of the pipe could be looked at as a mirror reflector of the waves, and the waves themselves as long light waves.

It may be recalled that, in the transmission-line-type wave, the wave must travel with the velocity of light, and the field-distribution pattern over the cross section must be identical with that which would have existed over that cross section if the line were energized by a d.c. voltage or d.c. current. Now, inside a very long hollow cylinder no electrostatic field can exist. Consider the pipe of Fig. 89. There can be no static electric field inside because

there is no way of maintaining it. To be sure, some electric field can be injected at one end, as the probe tries to do in Fig. 89, but that would turn out to be simply an end effect. The field would rapidly die out with distance from the source so that, except for the end effect, no field would exist.

But another way of putting the point may advance the argument more quickly. The chief interest for the moment is that no electric field at all, not even one diminishing with distance, can exist across the cross section exclusively transverse to the axis in the electrostatic case. Figure 90 shows a cross section of the guide of Fig. 89. If static electric-field lines are to lie in this plane, they must begin and end on charges. The lines would then have to go from positive charges on one part of the conductor's cross section to negative charges on another.

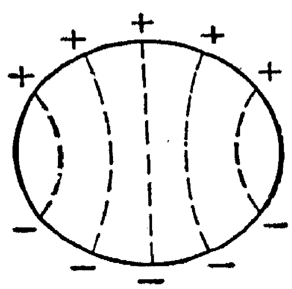

FIG. 90.

But this would be impossible in the electrostatic case. The charges would run together immediately, and the field would disappear. The conductor must be at one potential when no time variations are permitted.

Now all that this means is that transmission-line-type waves cannot be expected to propagate down the inside of a conducting pipe. Such a wave, with its static cross-sectional distribution and its insistence upon the velocity of light in free space, is, after all, a special wave. The fact that this particular wave is ruled out in a hollow pipe is unfortunate, perhaps, but it represented only one of the possibilities. Are there not other waves which can be sent down a tube?

Rectangular Wave Guide.—Next, examine one wave that may be made to travel down a rectangular wave guide (Fig. 91). (Whether a circular or rectangular cross section is chosen depends only on convenience.) First of all, imagine that the two sides A and B on the drawing are absent. Then there remains simply an ordinary two-

conductor transmission line in which the conductors happen to be wide, flat plates. When electric field is applied between top and bottom by use of a simple voltage generator, as shown in the figure, the ordinary transmission-line-type wave will propagate between the lines. Between

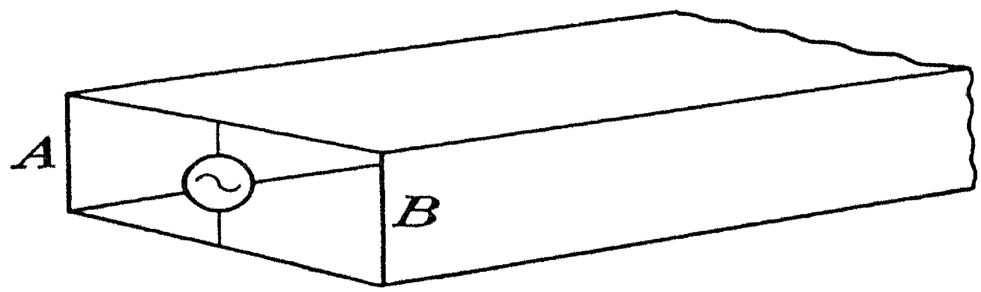

FIG. 91.—A wave guide of rectangular cross section.

the two conductors, especially along the center lines of each, there will be a high electric field. The electric-field distribution over the cross section would look something like that pictured in Fig. 92. This is simply a static distribution, which has been found to be completely justifiable.

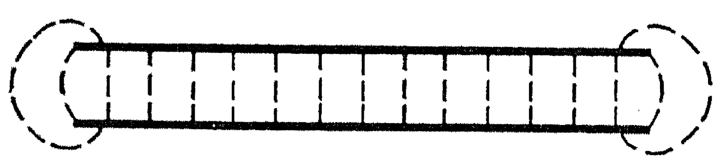

FIG. 92.

Now imagine that the two sides A and B are added again, but this time that A and B are not two short-circuiting bars made of good conductor. Instead, A and B will be high impedances. They can be resistive or reactive, but think first simply of their magnitudes as being high. The

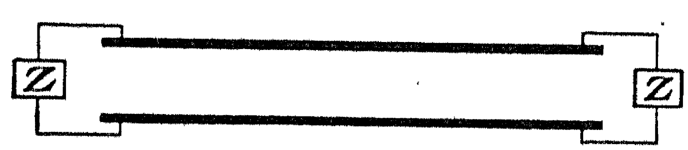

FIG. 93.

picture of Fig. 93 is then applicable for a cross-sectional diagram. With the Z of Fig. 93 high enough, the operation of the line is essentially unchanged.

Since Z can be realized by reactance, one way to obtain it is indicated in Fig. 94, which shows a two-plate trans-

mission line with a distributed single-turn inductance on each side. If the gaps *GG*, connecting the inductances with the line, are closed, the wave guide of Fig. 91 again results. It cannot be said from this that the rectangular guide, excited by a voltage between top and bottom plates, is the equivalent of the flat-plate transmission line and that, despite previous conclusions, it passes an ordinary transmission-line-type wave. The reactance that is distributed in parallel with the line is not infinite, so it is to be expected that the wave guided by the rectangular guide will be altered from the true simple line wave by virtue of the effect of the reactance. The reactance can be anything from zero to infinity; therefore it appears that it can do

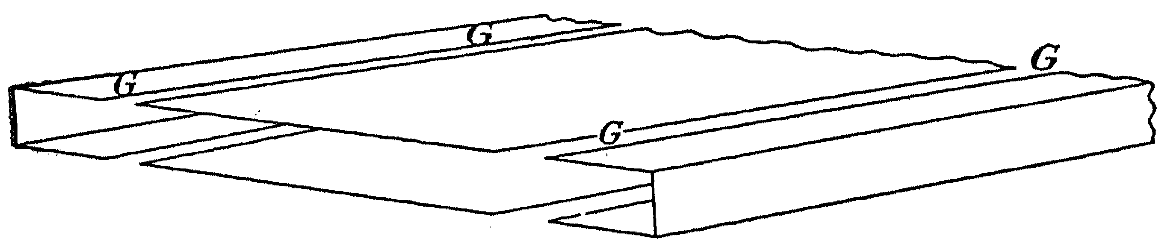

FIG. 94.—A distributed single-turn inductance is applied in parallel with the conductors of the line.

almost anything—from shorting the line, and preventing any wave propagation, to the other extreme of having no effect and allowing the ordinary transmission-line wave to propagate.

Apparently, it must next be known how large a reactance has been distributed along the two-plate line to form the rectangular wave guide. Then, with its magnitude known, it will be possible to see how the addition of such a reactance helps to set the distribution of fields in the guide, the wavelength and velocity, and other characteristics of waves guided by a pipe. If the section of the guide labeled a "distributed reactance" is to be studied for its reactance properties, something must be known of the current and magnetic-field distributions. No attempt will be made simply to look up the inductance of a U-shaped conductor in some table of inductances, for the criticism could then be made that the very meaning of the inductance is

doubtful for a long conductor along which current density may vary greatly with distance.

Instead, the discussion of fundamentals is resumed. To ask if the side or parallel distributed system looks like an impedance to the applied voltage generator (Fig. 95) is to ask whether it is able to tolerate a zero electric field at

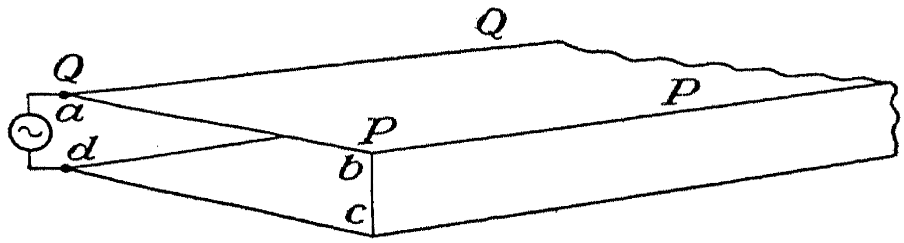

FIG. 95.

the edge PP while still possessing the required nonzero electric field set by the generator along the center line QQ. If there is electric field at QQ, none at PP, and none, of course, tangent to the surface on the perfect conductor, then around a closed path such as *abcda,* a net induced e.m.f. must exist. If so, a changing magnetic flux must exist linking this path. The first conclusion, therefore,

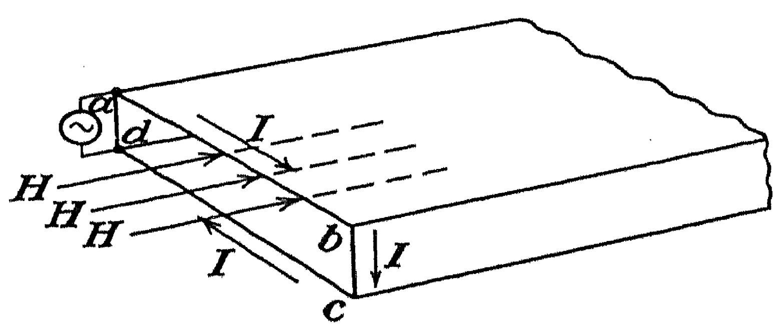

FIG. 96.—To maintain the cross-sectional a.c. voltage difference, some longitudinal a.c. magnetic field H and transverse side current I must exist.

is that there is some longitudinal or axial magnetic field in the guide as pictured in Fig. 96. Accordingly, some side current will flow around the loop *abcd* to produce this magnetic field.

If the operating frequency is high enough, only a slight side current need flow, producing a slight axial magnetic field. The high *time rate of change* of magnetic flux linked by that side current will be sufficient to support the difference in electric field between the center and the shorted

side of the wave guide. Obviously, as the frequency approaches infinity, the amount of current that must flow transverse to the direction of wave propagation will approach zero. The effective impedance at infinite frequency from the center line QQ toward PP is infinite because no current flows out sidewise. The wave can proceed along the center as though the shorting bars were absent.

Field Distributions.—Figure 97 shows a distribution of instantaneous electric field over the cross section of the rectangular guide that is consistent with the above discussion. The electric field is higher at the center; it falls off to zero at the sides. The axial, or longitudinal, magnetic flux, changing with time, is responsible for

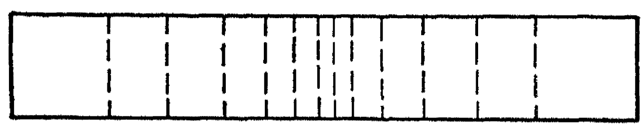

FIG. 97.

this dropping off of field. The higher the frequency the more does the wave act like a transmission-line-type wave. In other words, the same ratio of axial current to voltage occurs over the central portion of the guide as for two separate flat plates; the same variation of voltage and current with distance occurs; the same velocity of propagation and the same wavelength are obtained. But there is at least one difference even at the highest frequency between this wave for the pipe and the ordinary wave for the two-plate line: The voltage between top and bottom of the pipe (integrated transverse electric field) varies over the cross section. For a transmission-line wave on the two-plate line, the two plates were spoken of as having a certain definite voltage difference at each point along the line. In the hollow-cylinder guide, that voltage difference is one thing at the center and another thing (zero) at the shorted edges.

This seems to say, quite properly, that the periphery of the rectangular guide (speaking of the cross section only)

is not at one common voltage. If the idea of a piece of conductor's being at different voltages all at one instant is confusing, recall that this same condition exists in the ordinary line wave. As Fig. 98 shows, in this simple case the voltage difference between lines varies with distance along the line at every instant. Of course, here the variation happens to be in a different direction, *viz.*, in the direction of propagation. In the pipe wave guide, the

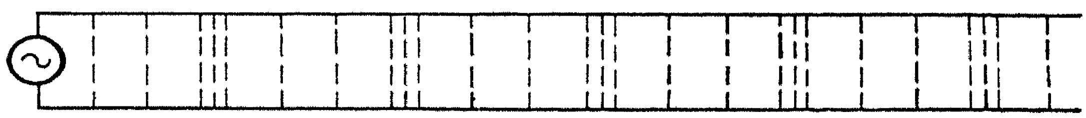

FIG. 98.

variation of voltage takes place in two directions simultaneously. It is a two-dimensional transmission line in this sense.

Strictly speaking, this idea that a conductor must be at one voltage is an idea that is usually learned first in statics. It happens to hold true over the active cross-sectional plane of the ordinary uniform line wave but not in the propagational direction even for that one special wave.

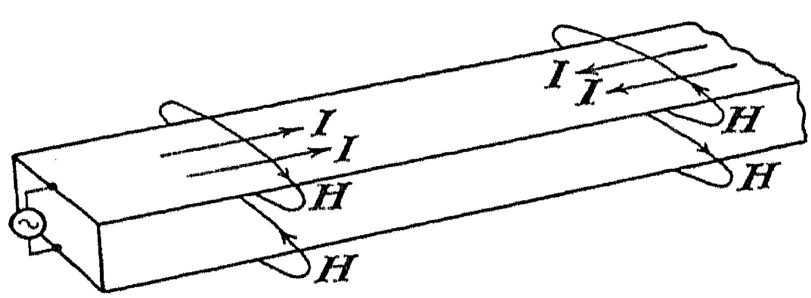

FIG. 99.—A few of the magnetic field lines H in the transmission-line-type wave at one instant.

It does not hold true for either the transverse, or axial, plane of a hollow cylinder passing a wave-guide-type wave.

Next, a little further study may be given to the distribution of magnetic field in the hollow-cylinder wave guide. It has already been noted that for the example chosen there must be axial field. There are other differences between the wave-guide wave's magnetic field and the transmission-line wave's magnetic field. Figure 99

shows the instantaneous magnetic-field distribution in the two-plate transmission line. The magnetic field is at all times transverse to the axis of propagation. The flux lines link the current, of course, and form closed loops.

When the shorting sides are added, the skin effect will prevent the flux lines from going outside and around the

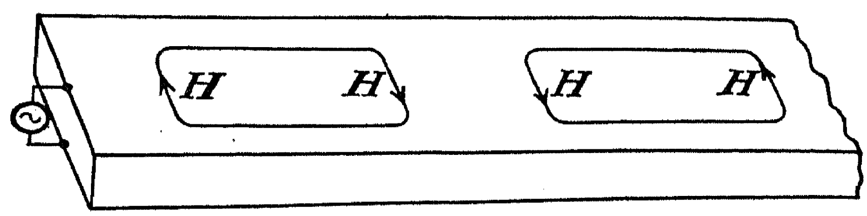

FIG. 100.—Two of the magnetic field lines H on the inside of the rectangular wave guide.

plates to close the path. Also, as has been deduced, some axial magnetic field must now exist if a wave is to exist inside the newly formed pipe. These two requirements cooperate very neatly to yield the instantaneous field pattern of Fig. 100.

Frequency Characteristics.—Figure 100 and the preceding diagrams and discussion have fairly well justified the idea of thinking of the hollow-cylinder wave in terms of

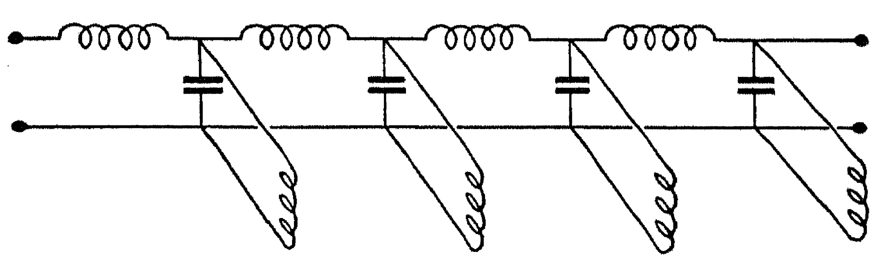
FIG. 101.—A representation of a wave guide carrying one type of wave.

distributed shunt reactance added to the ordinary line wave, which is represented so well by distributed series and shunt reactance. It seems now completely safe to look at the wave that has been studied as depictable by the circuit diagram of Fig. 101. The added shunt inductance represents the effect of the axial magnetic field, or the side currents. Notice that as the frequency becomes very high, the newly added shunt inductance has negligible effect. If the frequency is very low, the shunt inductance

steals all the input current and effectively shorts the line. In fact, the shunt inductance, at some sufficiently low frequency, so overshadows the shunt capacitance in the amount of current it takes that the diagram might be redrawn as shown in Fig. 102, where the new shunt induct-

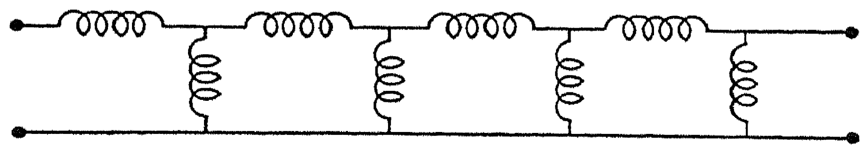

FIG. 102.—The wave-guide representation at low frequency.

ance yields the equivalent reactance of the true shunt inductance and capacitance.

Figure 102 clearly indicates how no propagating waves are possible if the frequency is too low. At some sufficiently low frequency, known as "cutoff frequency," the shunt reactance becomes inductive. The system then is

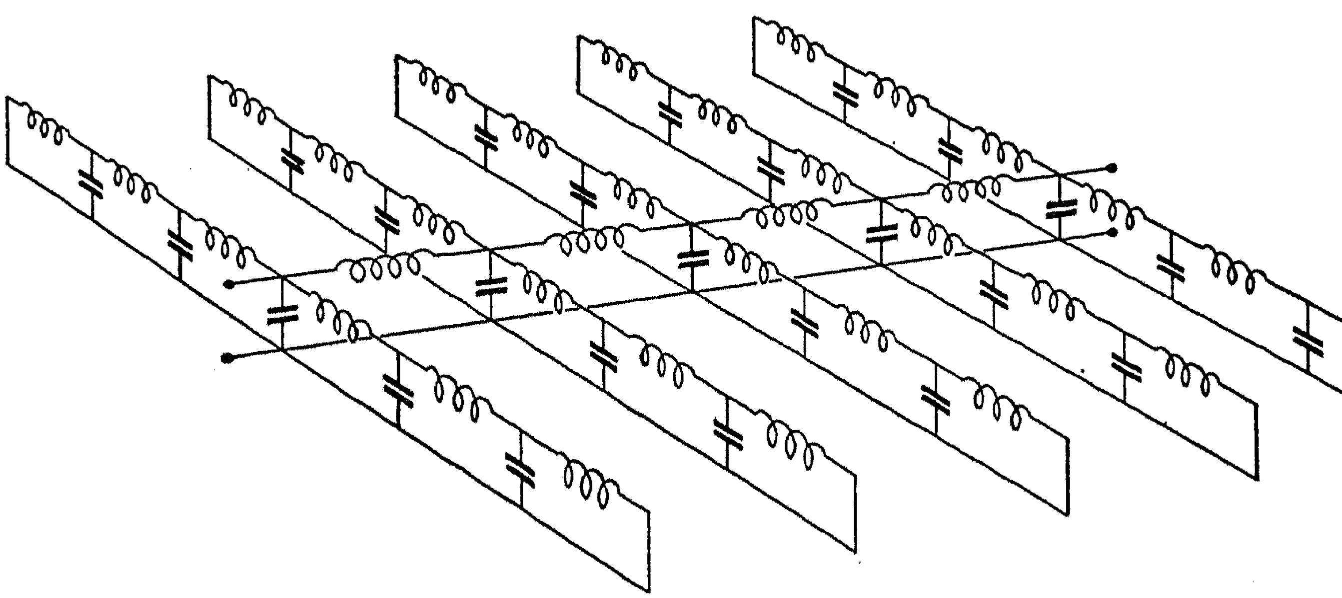

FIG. 103.—A better representation of the wave-guide-type wave that discloses the cross-sectional as well as longitudinal variations.

just an attenuator. All phenomena are in phase as distance is changed, since the impedances are all inductive reactances. Only the magnitude falls off with distance.

Of course, both Figs. 101 and 102 fail to show the cross-sectional variations. Figure 103 is more completely descriptive of the wave that has been studied. Completeness obviously does not alter the foregoing remarks,

but only offers the opportunity to see the cross-sectional variations also in terms of a distributed-constant network.

Wave Velocity.—Many new wave concepts have been uncovered in studying just one wave that can exist in a hollow cylinder. One point that is certainly worth a little more discussion is the fact that not all waves travel with the velocity of light—even if discussion is restricted to perfect conductors. In the simple transmission-line-type wave, the velocity is always that of light. It is, in other words, a constant independent of frequency. If frequency is lowered for such a wave, the wave compensates for it by changing its wavelength just enough to keep the velocity the same. Figure 104 shows how the voltage between

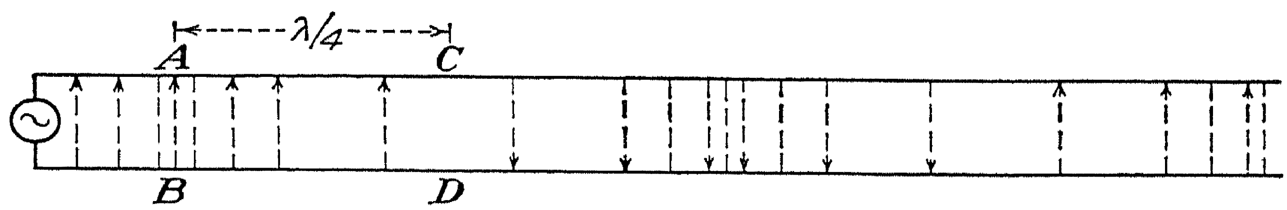

FIG. 104.—With the same magnitude of voltage, the rate of change of magnetic flux between a zero and maximum must be independent of frequency.

lines must fall with distance along the line from maximum to zero in a quarter wavelength. If the frequency of the generator shown were lowered, its voltage remaining the same, the wave would be forced to provide the same rate of change of magnetic flux linking the closed path *ABDCA* —this despite the lowered frequency. The wave does so, not by drawing more current to build up the flux density, but rather by extending the length $\lambda/4$, where λ is the wavelength. Doing this in exact proportion to frequency keeps the flux linked exactly right and holds the velocity a constant independent of frequency.

The common (hollow-cylinder) wave guide cannot do this. It must worry about the side currents. If the frequency is lowered, the wave tries to extend the wavelength to compensate just as the ordinary line did; but the side currents insist upon increasing for reasons discussed previously. (The side reactance is less at the lower fre-

quency.) This makes it even more difficult for the wave to link enough flux, so it must extend the wavelength even more, more than enough to keep the velocity constant.

At the extremely high frequency, with the side currents small, the wave holds its own as frequency is lowered and the velocity stays close to that of light. But as the critical cutoff frequency approaches, the wavelength overshoots in order to maintain itself, since the distributed shunt parasite is calling for more and more current. The velocity becomes higher as the wavelength becomes longer. Finally, in a desperate effort to preserve itself at the critical cutoff frequency, the wavelength becomes infinite. Everything

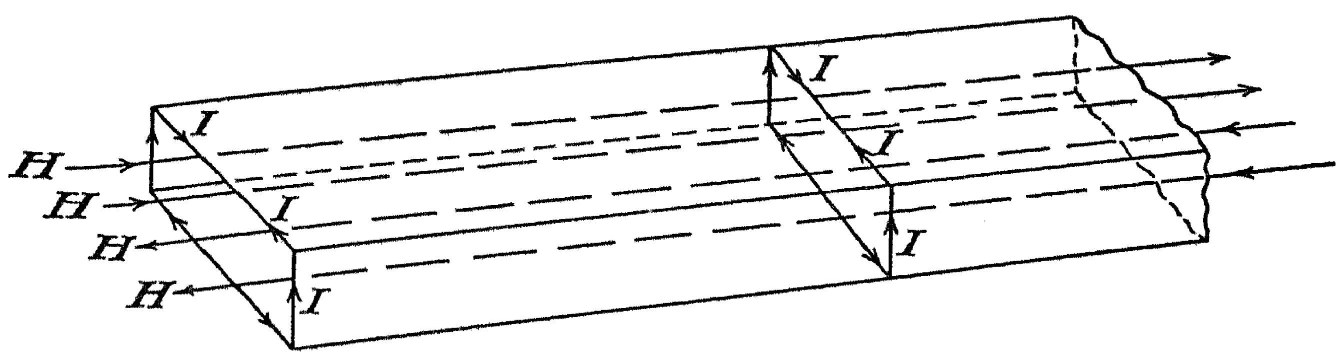

FIG. 105.—At the cut-off frequency, all the current I is transverse. The magnetic and electric fields and the current have phase angles that are independent of distance along the guide; the phase velocity is infinite.

in the guide is of like phase and the velocity is infinite. Figure 105 shows this for the wave studied here. The magnetic field goes right down the axis in constant phase and magnitude and is linked by the cross-sectional current, including the displacement current and the conduction current on the walls.

This condition, a critical one, is nevertheless in complete agreement with the fundamental laws. The velocity of the wave, as the term has been used throughout this text, is usually called a "phase" velocity. It simply equals the product of wavelength and frequency and is not equal the velocity of light in free space except in certain special instances. They happen to be common instances, but other waves are possible and practical for which the phase

velocity may be quite different from that of light. Similarly, when someone speaks of "50-cm wavelength" or "1-m waves," etc., he usually refers to the simple plane wave, free-space wavelength. For every frequency, the phase velocity and the wavelength are functions of the boundaries and the type of wave being guided by those boundaries.

Only one type of wave that can exist in a hollow cylinder has been considered. Many other types of waves can propagate inside a pipe. All have a critical cutoff frequency that depends upon the cross-sectional dimensions and shape. All have regions in the frequency spectrum where there is propagation. All have phase velocities and wavelengths and field distributions that depend not only upon frequency but also upon the cross-sectional geometry. All of them have characteristics that approach the ordinary transmission-line-type wave as the frequency approaches infinity. All are characterized by having either axial electric fields or axial magnetic fields, whereas the ordinary transmission-line-type wave has only transverse or cross-sectional electric and magnetic fields. Sometimes the waves are labeled *H* or *E* waves to denote the presence of axial magnetic or electric field. A more common notation is to label waves *TE* or *TM* for "transverse electric" or "transverse magnetic," to indicate modes in which the electric or magnetic field is entirely transverse to the axis of propagation. All have cross sections that contain certain dimensions comparable with wavelength. Either the span of the cross section, the diameter, the distance between top and bottom, or the inner perimeter of the conductor is comparable with wavelength.

For all these waves it is true that the propagation can happen at any frequency. This is an example of one point raised in the early chapters: There is nothing that happens at ultra-high frequency that could not happen with physically larger systems at low frequency. It is the ratio of the cross-sectional dimensions to wavelength that is

important; if the cylinder is made large enough, then propagation can take place at any frequency.

All this raises the question of how the electromagnetic wave can decide what form it will take, what kind it shall

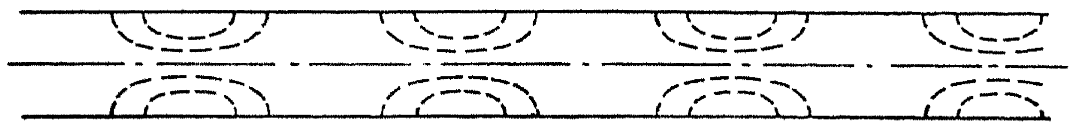

FIG. 106.

be for any given system. For a hollow, circular cylinder, should it be one such as is shown in Fig. 106 or should it be as shown in Fig. 107? This, it will be seen, is almost a separate subject in itself; it deals with the ideas of getting waves started. It can only be said at this point that the

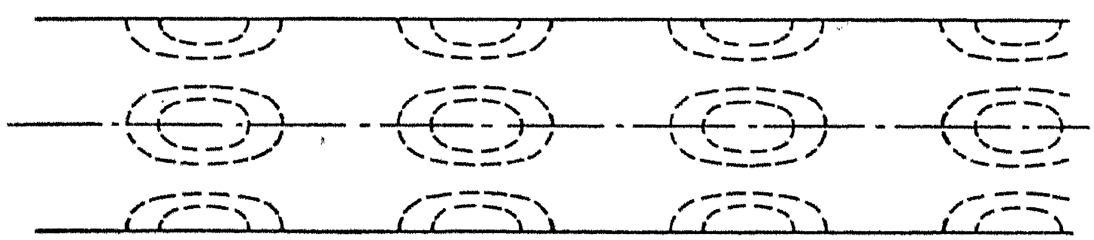

FIG. 107.

starting mechanism, which is an end effect, is very important. A wave has to be started before it can propagate. Often it happens that many waves are started, but most of them attenuate rapidly, since for them the cross section is not yet large enough or, in other words, the frequency is

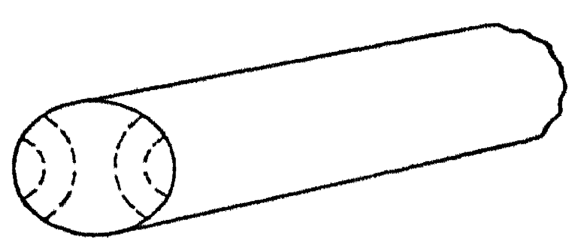

FIG. 108.

not yet high enough. For other waves the frequency is sufficiently high to allow propagation.

Figures 108 and 109 indicate two other distributions of electric fields that can take place in wave guides. Other distributions can be drawn, as in the case of the resonant

cavity. The distribution of the magnetic field and the electric field can always be understood by simply holding to the basic notions: displacement currents act in con-

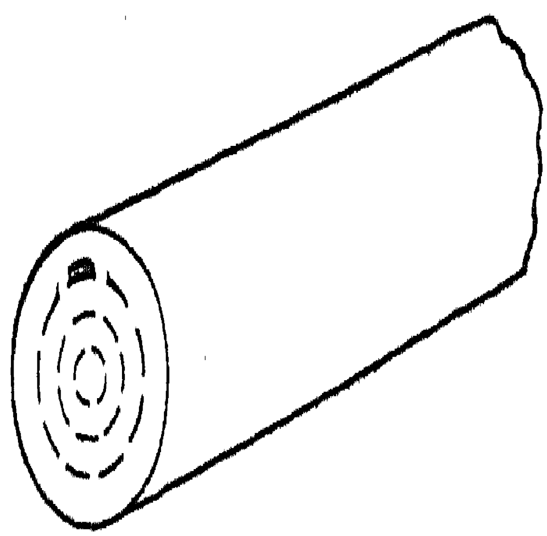

FIG. 109.

junction with the conduction current on the walls to form closed systems that link with the magnetic-flux lines; the magnetic flux lines close to link changing electric field.

CHAPTER XIII

Microwave Phenomena May Be Thought Of as the Result of Combining a Series of Various Waves

Wave types discussed so far are the simple plane waves common in free space and on transmission lines and the waves that will propagate down the inside of a hollow conducting cylinder. There is nothing new in concept to say that there are many more wave types that correspond to different guiding boundaries and different starting mechanisms. But perhaps it is a broadening of concept to point out the importance of being able to visualize any microwave manifestations as the result of combining a whole series of different waves.

In an exceedingly simple transmission-line problem there may be only one wave to worry about. In a more complex case there may be an improper termination at the end of the line, and a reflected wave may have to be considered in addition to the incident wave. But even so simple a situation as this is not as common as might be wished at microwave frequencies. For these high frequencies, the uniformity of lines and guides is disturbed by relatively huge end effects. The connections between lines and wave guides and electronic tubes and resonant cavities have such complex geometrical configurations that hosts of waves that are not the principal ones are started everywhere.

This is not just a minor, secondary effect. It is everyday microwave life. A centimeter-wave engineer learns to live with these many series of different waves—to separate them, evaluate them, and see how they add up to give the resultant voltage and current effects at the important terminals.

Consider the simple change of cross section in a coaxial line as pictured in Fig. 110. It is easy to see how the single break in uniformity of the line immediately causes the more complex waves to be present in the line in addition to the simple transmission-line-type waves which the lines may be propagating. Certainly, it might be expected that a simple wave traveling from line A to B would be partially reflected and partially transmitted, and this is, indeed, found to be true. But at the higher frequencies this problem cannot be dismissed by studying the combination of these two or three simple waves.

The simple waves cannot be added in any combination to satisfy the boundary requirement along the conductor

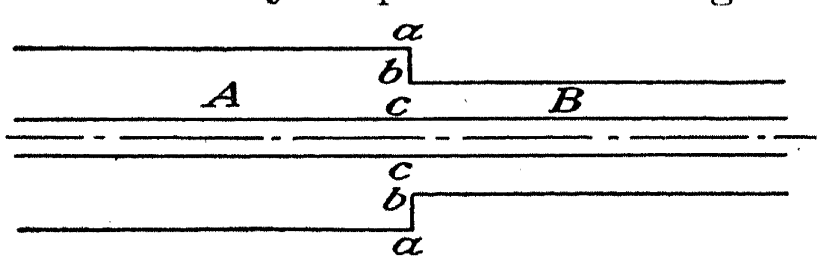

FIG. 110.—Higher order waves are started at all discontinuities in general.

in the radial region ab (Fig. 110). The simple waves always have transverse (in this instance, radial) electric field; the conductor ab will not tolerate any radial electric field since the radial direction is tangential to its surface. The only way the conductor in the region ab can be pleased by simple waves alone is to have the electric field equal to zero over the whole cross section abc. This is a situation in which no wave energy would cross the discontinuity and is certainly a very special case.

In general, the surface ab will protect itself against the tangential field of the simple principal waves by insisting upon the presence of other more complex, so-called "higher order" waves. These higher order waves will have just enough amplitude to buck out any unallowable boundary effects of the simple principal waves that are generally present alone only for a completely uniform system. The higher order waves, like the wave-guide waves of Chap. XII, have critical frequency characteristics. They will

often die out quickly; they are usually localized around the discontinuity. But at the break the additional waves figure in the determination of the ratio of reflected to transmitted waves, the distribution of fields, and the distribution of voltage and current.

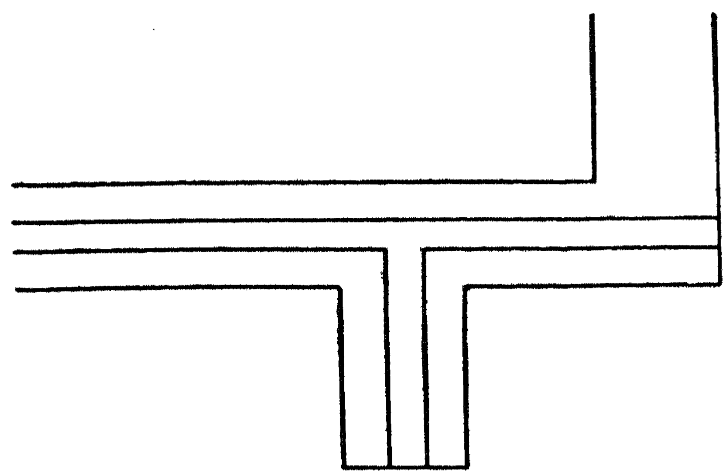

FIG. 111.—Typical microwave systems consist of much discontinuous plumbing.

Since a typical microwave system may consist of such combinations of plumbing as Fig. 111 depicts, it can easily be imagined how many different kinds of waves may have to be considered in order to dissect the effects contained therein.

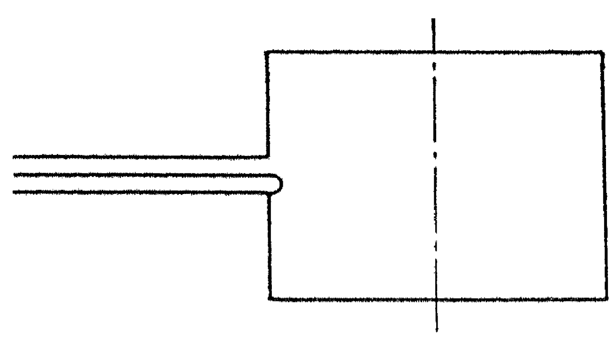

FIG. 112.—The current in the loop excites various modes in the cavity; the mode near resonance will have greatest magnitude.

Consider another example (Fig. 112). A hollow-cylinder resonant cavity is stimulated by current in a small coupling loop. The current in this loop starts azimuthal, or circular, magnetic flux. But there are many modes or field patterns that have azimuthal magnetic flux and that will be tolerated by the cylinder. Obviously, all these waves will add up with exactly the right amplitudes to the field called for by the loop. Again, the loop will react to the waves

it has started. It must defend itself against the electric fields that accompany the resonant cavity's waves and that insist upon having components tangent to the conductor of the loop.

The current flow in the loop for a given voltage applied to the line that feeds it will depend upon the series of waves started in the box. The input impedance to the loop will be different from its value in free space or in a different-sized cylinder.

This is a good time to introduce an analogy that pigeonholes the importance of this series-of-waves idea. Consider a simple circuit at low frequency that is excited by a square wave, or any nonsinusoidal voltage. The current is then thought of as the resultant of a fundamental and a series of harmonics. It is quite customary to think of nonsinusoidal effects in time as consisting of a series of sine waves. That technique is not necessary; it is simply convenient. It is entirely conceptual and we like it because it is so easy to think of sine waves and make calculations with them. In the simple-circuit instance cited, the voltage is broken up into an equivalent series of sinusoidal voltages, and the sine-wave current due to each component is found. It is both a way of doing the calculation and a way of thinking of what goes on. We conceive of all the impedance drops due to every current component summing to the applied voltage.

Space Harmonics.—In microwave systems, we have to conceive of the space phenomena as a series of space harmonics. It is not possible simply to apply nonsinusoidal time variations and deal exclusively with time-harmonic series. We must add the concept of applying space boundary effects and requirements and must deal also with the resultant space harmonics. The discussion cannot go further here, but the reader should appreciate why microwave engineers must be continually classifying and studying the many space distributions, the wave types of which all microwave phenomena may be conceived of as formed.

CHAPTER XIV

Voltage, Current, and Impedance Concepts Have Practical Use at the Highest Frequencies

The commonest concepts of low-frequency electricity are voltage, current, and impedance. Some, it is true, may concern themselves more with electric and magnetic fields. A few engineers or physicists must engage in setting up the self and mutual capacitances and inductances by use of field theory, so the rest can be content with working with circuit theory where the question of fields need hardly arise. At ultra-high and microwave frequencies it seems necessary to work intimately with the electric and magnetic fields for almost any part of almost any problem. It is true that the student who goes deeply into microwaves must develop a thorough understanding of field and wave theory. But in doing so, it is also important to locate each element of field theory with respect to over-all boundary effects, which are usually most conveniently stated in terms of voltage and current.

The text has sought not to overlook this latter point, but the issue has often arisen as incidental to some other main theme. It is well now to emphasize certain key ideas that will help to visualize the change or broadening of applicability of voltage, current, and impedance concepts as the frequency goes from low to very high values.

Voltage Concept.—Consider first the notion of voltage. It is useful of course at the highest frequencies. Does it mean the same thing for microwaves as for low frequencies? That depends upon the meaning of voltage accepted at any frequency. It is possible to hold to a conception of voltage that will be true and constant and useful over the whole

band. It is also possible to have notions of voltage that are good only for statics or for low-frequency circuits.

The idea of a conductor's being at one potential is perfectly all right for statics and generally all right for low frequencies; but at high frequency it is unsatisfactory and often completely erroneous. A safe and useful concept is to think of a voltage difference between two points as an integration of electric field between those two points along the electric-field lines. For statics, where there is no question of changing magnetic flux causing induced voltages, it is permissible and desirable to include as an additional concept that a conductor is an equipotential surface.

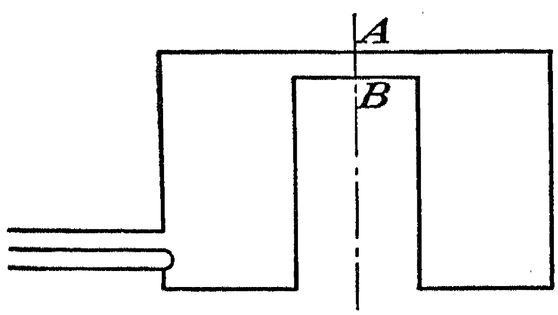

FIG. 113.

Because of skin effect, this notion of a conductor's being at one voltage with respect to another requires even greater caution in application. Consider the excited resonant cavity shown in Fig. 113. Here it is possible to integrate the electric field between the surfaces A and B and fix a voltage difference between these points even though they are on the same conductor. With the source on the inside, skin effect restricts all electromagnetic phenomena to the inside of the cavity, and an indicator on the outside of the cavity will show essentially no response in any practical case. Between A and B there is a voltage difference on the inside of the cavity, but essentially no electric field exists between them on the outside. Before, when it was said that the voltage difference could vary between two conductors, variations with distances appreciable compared with wavelength were meant. Here the voltage of a conductor appears to vary radically in shifting the point of investigation a few hundredths of an inch. The conclusions are correct. If they conflict with any notions about how the term "voltage" should be used in connection with con-

ductors, the conflict is a further indication that such notions should be tied closely to the electric field.

In general, voltage is an over-all effect, an integration result—not in the mathematical sense particularly but in the physical sense. Its successful application, its legitimacy, can be verified by a study of the field distribution. This is true at any frequency. In a transmission-line-type wave the voltage-difference idea is completely applicable. In a general transmission-line problem, where there may be discontinuities that will cause higher order waves, the voltage idea may be applied liberally nevertheless, at least to the principal simple waves.

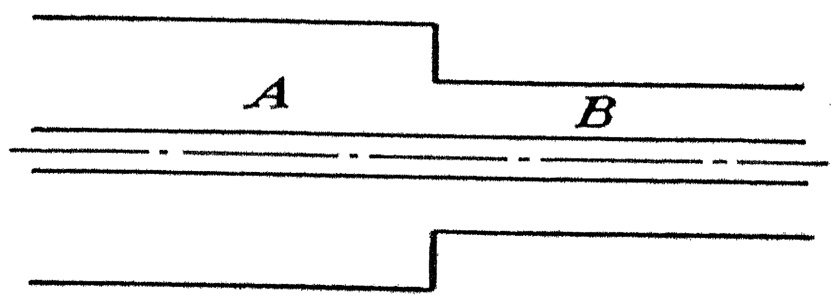

FIG. 114.

For example, consider again the coaxial-line discontinuity shown in Fig. 114, and assume that the distributed L and C or the equivalent has been previously determined. It is not necessary in the case of the principal waves sent up and down the lines, to worry about fields but only about the over-all voltages and currents. For the higher order waves, the existing field problem must again be recognized. However, one single skillful handling of that field problem makes it possible for us to forget it in application by finding out how much the voltage and current of the principal waves vary because of the presence of the higher order waves.

Such an outflanking movement leads to an equivalent circuit for the situation of Fig. 114 as pictured in Fig. 115. Here the two original lines are connected, but a lumped capacitance is placed across the line at the discontinuity.

The rules are to integrate the effects of all the field disturbance into an over-all effect on the voltage and current of the principal waves. To know what happens when two lines are connected together, the fields at the discontinuity should be studied, like the fields in the uniform part of the line, and the over-all answer obtained in terms of a lumped

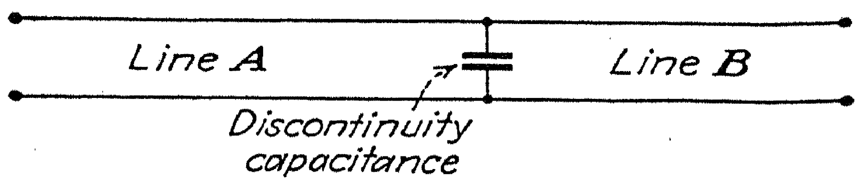

FIG. 115.—The field problem of the discontinuity is solved in the convenient form of a lumped impedance.

capacitance. If only the over-all voltage and current of the system are of interest, the problem then proceeds with only the voltages and currents of the uniform-line-type wave entering the analysis. The equivalent circuit of Fig. 115 will not give the distribution of things around the discontinuity—but neither will the distributed LC approach to a uniform line give the field distribution over the cross section, nor the lumped circuits of low frequency give the magnetic-field strength in the circuit's vicinity.

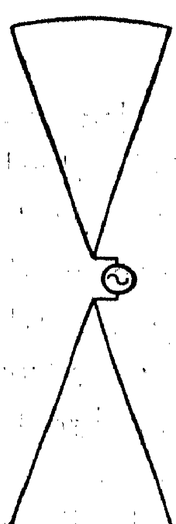

FIG. 116.

Every indication is that, no matter how complex the fields become, there are always important parts of the problem where it is desirable to define and use the over-all voltage effect. Consider the symmetrical antenna of Fig. 116. Here waves start out with every intention of being simple (Fig. 117), the electric field following circular arcs and the current going out radially. Again, it is useful and possible to think of a voltage between cones, the integrated electric field. Here the discontinuity is severe when the outgoing radial wave finds the line has ended. The smooth line ends abruptly, and a series of waves is started. That, of course, is the intentional function of the system—to radiate waves out to the space. But one of the things an engineer wants to know about the antenna

is its input impedance, the ratio of the voltage to current at the entrance terminals of the antenna. Here again, in an antenna, it is correct and useful to accumulate all the effects of the waves that result from the discontinuity into a lumped impedance presented to the simple outgoing wave—the wave that would have existed alone had the cones extended out to infinity. The field problem must be solved to find the value of the impedance to lump at the end. But it is a good concept to think of the lumped impedance and the simple wave's voltage and currents as a representation of some aspects of the problem.

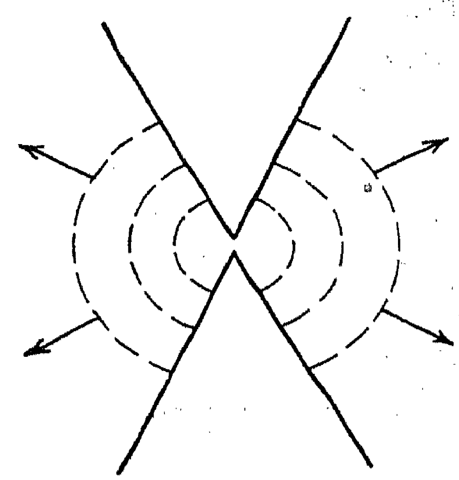

FIG. 117.—The outgoing principal wave on the antenna of Fig. 116.

Current Concept.—In the discussion of voltage, current and impedance have already been brought in. The current concept probably requires less justification than the voltage concept. In fact, there is no need to justify it: the electric charge and its motion are a perfectly good beginning point on which other concepts, even the fields, can be based. The matter that needs summarizing here is rather that, in studying microwaves, the current concept must include: (1) current sheets, that flow on surfaces of conductors and that bound the waves, keeping them from penetrating the surface; (2) induced current on conductors, due to motion of free electric charges in the surrounding space; (3) displacement current that is indistinguishable from conduction or convection current in producing magnetic field.

Impedance Concept.—The impedance concept finds obvious application wherever the voltage and current concepts do. Thus to a great extent impedance has already been discussed. In addition to its applicability as a ratio of voltage to current, the impedance concept is of important use in field and wave problems where the distribution of effects, not merely over-all quantities, has to be studied.

Microwave engineers are coming to think of impedance in electric phenomena not only as a ratio of voltage to current but more broadly as a ratio of electric to magnetic field. This impedance can be attached to a point in space just as can electric- and magnetic-field intensities. Also, since any point in space may have many different ratios of electric- to magnetic-field amplitudes, depending for one thing on the type of wave being transmitted, every point in space may have different *wave impedances*.

It is possible to describe wave types in terms of wave impedances. The convenience of this concept cannot be made entirely clear without going further into the matter of solving problems than space permits in this text. But some easily mentioned viewpoints will make it seem reasonable that the concept should be a valuable one.

Consider, for instance, the stress placed upon the discontinuities that cause many complex waves and upon the importance of such higher order waves. Since the impedances of a wave are a point-by-point description of it, the question of whether waves will be able to cross boundaries, merge and form into new waves, or be reflected can all be put on an impedance basis. One space can be thought of as matching another, just as impedances match in ordinary circuit theory. Through the application of impedance notions to fields and waves as well as to lumped and distributed circuits, many ideas from circuit and transmission line can be transferred immediately to waves. For example, a wave can pass smoothly from one medium into another, crossing the boundary of separation between them without any reflection if the wave impedances of the two media are the same.

Voltage, current, and impedance—concepts coming from d.c., or low frequency—seem destined to stay in the picture at ultra-high frequency. The problem is to recognize how they must be expanded or altered, and especially how fields and waves must be merged with them to form a completely solid and practical picture of electricity at any frequency.

CHAPTER XV

A Microwave Radiating System Combines Concepts from D.C. to Light-wave Frequencies

There are several reasons why some remarks on antennas are proper in a discussion of microwave concepts. First, the ease with which waves can escape from an unenclosed microwave system makes all microwave devices potentially akin to antennas. After the study of antennas, microwaves are better understood.

An antenna is, of course, not exclusively an ultra-high-frequency device; it is just as valuable a component of lower frequency radio systems. But, to be successful as a radiator, its dimensions must be (or preferably should be) made comparable with wavelength. At any frequency, then, a good antenna is illustrative of microwave phenomena, for it is *dimensions compared with wavelength* that counts, not the absolute frequency.

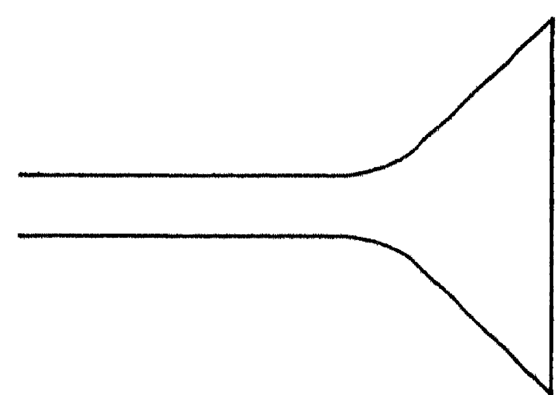

Fig. 118.

Microwave antennas are often different from low-frequency antennas because advantage is taken of the small wavelength to use concepts suggestive of light and sound waves to form directive systems.

Finally, it might be said that nothing demonstrates the merging of concepts stemming from direct current to light-wave frequencies better than a typical dipole microwave antenna in a parabolic reflector.

First, note the radiator of Fig. 118. It starts out as a wave guide and flares out gradually into a horn. Such a device, just as in the case of a horn that radiates sound

waves, can be made to transmit a well-focused or directed beam. The opening area should contain a large number of wavelengths if this is to be true. This becomes a practical thing to accomplish as wavelengths come down into the centimeters. It is theoretically possible to use directive horns for antennas at the lower frequencies, but the resulting structure would be impracticably large.

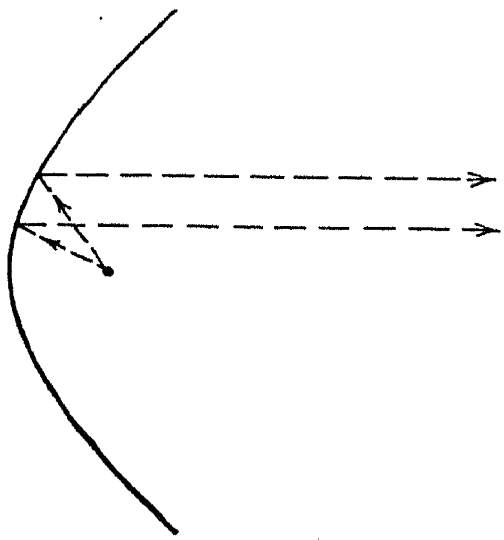

FIG. 119.—Light from a point source is reflected into a parallel ray beam.

Consider next the focused beam headlight of Fig. 119. This light system is described by conventional geometrical light-optics theory about as follows: A small source of light is placed at the focal point of a parabolic reflecting surface. The shape of this surface is such that the rays from the source are reflected in the direction parallel to the axis. Thus a highly directive beam is produced.

Since the difference between light and radio waves rests on the difference between their wavelengths, a similar directive system should be possible for radio waves if the wavelength becomes short enough. Microwaves are short enough to make these concepts useful for radiating systems.

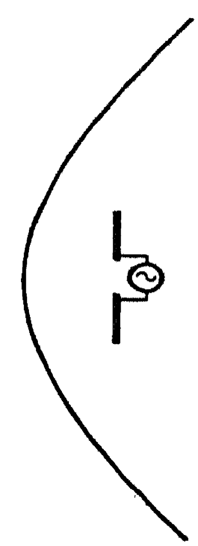

FIG. 120.—A dipole antenna in a parabolic reflector.

Figure 120 shows a so-called "dipole antenna" placed at the focus of a large parabolic reflector. At the input to the dipole, voltage is applied, and current flows up and down the two wires. As a result, waves are radiated. They strike the parabolic conductor and are reflected in the general axial direction more or less parallel, just as in the headlight.

In concept, at least, conventional voltage and current excitation applies, much as in the input to a d.c., or low-

frequency, circuit. However, a large part of the power passing into the antenna is not dissipated in ohmic losses but rather in radiation, in a transfer to radio waves in free space, a phenomenon ignored by low-frequency-circuit concepts. The directive action of the reflector is next deduced from the geometrical optics of light.

For real accuracy in this analysis, it is well to look to the basic wave equations. The dipole antenna is not a point source in the first place. For that reason alone the system should not be expected to send out an absolutely parallel beam. But a more fundamental consideration is that the use of geometrical optics is an approximation. It becomes a closer and closer approximation as the wavelength becomes shorter. In other words, it is an approximation that becomes better as ratio of the wavelength to the dimensions of the apparatus becomes smaller. It is an excellent approximation for most light systems; less so for microwaves.

The text has sought to encourage further microwave study by continually emphasizing the need for using either broad concepts or narrow ones whose limits are well understood. One of the problems throughout concerned the validity of ideas obtained from frequencies lower than microwave frequencies. Microwaves are so located in the frequency spectrum that there must also be some concern about adopting narrow ideas, like the rectilinear projections of geometrical light optics, from frequencies higher than those of microwaves.

APPENDIX

Textbooks for Further Study of Microwave Engineering or the Underlying Physics

EMERY, "Ultra-high-frequency Radio Engineering," Macmillan, 1944.

RAMO and WHINNERY, "Fields and Waves in Modern Radio," Wiley, 1944.

TERMAN, "Radio Engineers' Handbook," McGraw-Hill, 1943.

HARVEY, "High Frequency Thermionic Tubes," Wiley, 1943.

SARBACHER and EDSON, "Hyper and Ultrahigh Frequency Engineering," Wiley, 1943.

SCHELKUNOFF, "Electromagnetic Waves," Van Nostrand, 1943.

BRAINERD, KOEHLER, REICH, and WOODRUFF, "Ultra-high-frequency Techniques," Van Nostrand, 1942.

SKILLING, "Fundamentals of Electric Waves," Wiley, 1942.

SLATER, "Microwave Transmission," McGraw-Hill, 1942.

STRATTON, "Electromagnetic Theory," McGraw-Hill, 1941.

SMYTHE, "Static and Dynamic Electricity," McGraw-Hill, 1939.

Index